DOCUMENT RESUME

ED 101 511 95 EC 071 229

AUTHOR	Asp, Carl W.
TITLE	The Effectiveness of Low-Frequency Amplification and Filtered-Speech Testing for Preschool Deaf Children. Final Report.
INSTITUTION	Tennessee Univ., Knoxville.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
BUREAU NO	522113
PUB DATE	30 Nov 73
GRANT	OEG-0-9-522113-3339(032)
NOTE	89p.
EDRS PRICE	MF-$0.76 HC-$4.43 PLUS POSTAGE
DESCRIPTORS	Audio Equipment; *Auditory Training; Aurally Handicapped; Curriculum Guides; *Deaf; Exceptional Child Research; Listening Skills; *Preschool Education; Speech Improvement; *Speech Skills; *Speech Tests; Teaching Methods; Test Validity
IDENTIFIERS	*Verbo Tonal Method

ABSTRACT

During a 3-year period, 20 preschool deaf children were matched and given auditory training by the Verbo-tonal method using two different amplification systems (one which amplified from 200 to 5000 hertz and the other from 20 to 5000 hertz). There were three main goals: (1) to compare two different amplification systems to determine if the addition of frequencies in the 20 to 200 hertz range could facilitate the acquisition of speech perception and speech production skills in young deaf children; (2) to evaluate if severely impaired children could be trained auditorily by using aural/oral procedures such as the Verbo-tonal method; and (3) to evaluate filtered-speech testing for young deaf children. There was no significant difference between the speech reception/production scores of the two groups. Both groups showed significant improvement, indicating that the verbo-tonal method was effective in the auditory training of severely hearing-impaired children. The filtered speech testing produced similar detection thresholds to those of pure-tone audiometry. (Appendixes include a review of literature on the use of low-frequency amplification with the hearing impaired, a Verbo-tonal method materials list, and a curriculum guide for using the Verbo-tonal method.) (Author/LS)

FINAL REPORT

Project No. 522113
Grant No. OEG-0-9-522113-3339 (032)

The Effectiveness of Low-Frequency Amplification and Filtered-Speech Testing for Preschool Deaf Children

Carl W. Asp, Ph.D.

The University of Tennessee
Knoxville, Tennessee

November 30, 1973

The research reported herein was performed pursuant to OEG-0-9-522113-3339 (032) with the Bureau of Education for the Handicapped, U.S. Office of Education, Department of Health, Education, and Welfare. Grantees undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.

Department of Health, Education, and Welfare

U.S. Office of Education
Bureau of Education for the Handicapped

ABSTRACT

Twenty preschool deaf children were matched and given auditory training by the Verbo-tonal method using two different amplification systems. During the three-year period one group received amplification from 200 to 5000 Hz and the other received amplification from 20 to 5000 Hz. There were three main goals: 1) to compare two different amplification systems to determine if the addition of frequencies in the 20 to 200 Hz range could facilitate the acquisition of speech perception and speech production skills in young deaf children; 2) to evaluate if severely impaired children could be trained auditorily by using aural/oral procedures such as the Verbo-tonal method; and 3) to evaluate the usefulness and appropriateness of filtered-speech testing for young deaf children. Progress in auditory training was measured through evaluation of the children's speech production during test sessions following periods of training. The children's speech sounds were often unintelligible and conventional intelligibility measures could not be used. Thus, new measures were developed. There was no significant difference between the speech reception/production scores of the two groups. Both groups showed significant improvment, indicating that the Verbo-tonal method was effective in the auditory training of severely hearing-impaired children. The filtered speech testing, a promising diagnostic test for hearing-impaired adults, produced similar detection thresholds to those of pure-tone audiometry. The new measures developed in this project are useful in evaluating progress in speech production from the sub-intelligible to the intelligible level. These include: a similarity scale, a phonetic count, and a battery of acoustic measures.

TABLE OF CONTENTS

	Page
Abstract	ii
Introduction	1
Rationale	1
Method	3
A. Equipment	3
B. Subjects	6
C. Teachers and the Tester	10
D. Experimental Design	11
Results	14
A. Similarity Scale	14
B. Acoustic Measures	20
C. Phonetic Count	29
D. Intelligibility	30
E. Additional Measures and Analyses	32
1. Analysis of the Type of Response	32
2. Hearing Aid Evaluation	32
3. Half-harmonics	33
Filtered-Speech Testing	33
Discussion	38
Conclusions	41
Acknowledgments	42
Appendices	43
A. The Use of Low-Frequency Amplification with the Hearing-Impaired: A Review of Literature	44
B. The Electro-Acoustical Parameters of the Equipment Used in the Study	49
C-1. Period of Enrollment for the Warren Group	52
C-2. Period of Enrollment for the Suvag Group	53
D. A Discussion of the Individual Subjects	54
E. Descriptive Information on the Teachers Employed on the Project	57

Page

F. A Description of the Verbo-tonal Procedures Used on This Project ---------------------------- 58

G. The Training and Supervision of the Classroom Teachers -- 60

H. The Verbo-tonal Method Materials List ----------- 62

I. The Testing Procedure for Obtaining Speech Samples --- 67

J. Procedures for Obtaining Acoustic Measures From the Spectrograms -------------------------- 69

K. Curriculum Guide -------------------------------- 71

References --- 82

INTRODUCTION

The purpose of this study was to evaluate the effectiveness of low-frequency amplification for training preschool deaf children. Two matched groups of preschool children were given auditory training by the Verbo-tonal method using two different amplification systems - one with a frequency characteristic of 200 to 5000 Hz and another with amplification from 20 to 5000 Hz.

There were three main goals of the study: 1) to compare two different amplification systems to see if the addition of the frequencies in the 20 to 200 Hz range could facilitate the acquisition of auditory discrimination and speech production skills in young deaf children; 2) to evaluate if severely impaired children could be trained auditorily by using aural/oral procedures such as the Verbo-tonal method; and 3) to evaluate the usefulness and appropriateness of filtered-speech testing for young deaf children.

Progress in auditory training was measured through evaluation of the children's speech production during test sessions following each four-month period of training. The children's speech sounds were often unintelligible and standard intelligibility measures could not be used. Therefore, new testing procedures were developed which proved to be useful in evaluating the progress in speech production from sub-intelligible to intelligible levels.

RATIONALE

Most educators of the hearing impaired agree that the auditory sense is the most suitable perceptual modality by which the normal-hearing child learns speech and language (Perkins, 1971; Pickett, 1972), but some educators continue to oppose the use of the auditory channel as the primary avenue for habilitating the deaf child (Vernon, 1972). A frequently-cited argument against emphasizing auditory stimulation is that the hearing of most deaf children is deficient to the degree that prohibits aural habilitation. However, the percentage of the children with residual hearing is quite large. According to Huizing (1959) and Watson (1961), between 95 and 97% of the children enrolled in schools for the deaf have some measurable hearing, usually below 500 Hz. These statistics encourage attempts to use auditory training for changing measurable low-frequency hearing to "usable" hearing. To accomplish this task, it is necessary to identify the most appropriate frequency response for auditory training.

Previous studies, notably The Harvard Report (Davis, et al., 1947), have recommended a frequency response from 300 to 4000 Hz with sharp cutoffs below and above this range as being "the best choice for all ears." These investigators stressed that amplified low-pitched components of ambient noise or background speech mask the high-pitched components of speech. However, their recommendation was based on results obtained with hard-of-hearing adults, most of whom had acquired losses. The frequency response which is optimum for auditory training of prelingually deaf children may differ from that yielding the best discrimination scores in adults with acquired losses.

There are several arguments against low-frequency amplification. First, the long-term speech spectrum (French and Steinberg, 1947; Benson and Hirsh, 1953) and the spectra of vowels (Fletcher, 1929) show little energy below approximately 90 Hz. Second, most of the energy for ambient noise is concentrated in the low-frequency range. Third, the sensitivity of the auditory system is not as great in the low-frequency region. Thus, some educators would predict a poorer performance for hearing-impaired children who are trained with the unit that includes an extended low-frequency response In spite of these logical arguments against low-frequency amplification, several investigators (Ling, 1963, 1964a, 1964b, 1966; Briskey and Sinclair, 1966; Briskey, Garrison, Owsley, and Sinclair, 1967; Leckie and Ling, 1968) reported that low-frequency hearing aids appear to benefit the performance of some deaf children.

Tied to this controversy of low-frequency amplification is the problem of developing intelligible speech patterns. Some researchers (Hudgins and Numbers, 1942; Hudgins, 1946; John and Howarth, 1965; Levitt, 1971) have stressed that suprasegmental errors (intonation, rhythm, phrasing) have a negative effect on the intelligibility of deaf speech.

Guberina (1964), the originator of the Verbo-tonal method, contends that it is precisely these types of errors that can be alleviated by auditory training with extended low-frequency amplification. He maintains that low-frequency amplification is essential for developing normal intonational and rhythmical patterns in the speech of deaf children, and that the normal development of these suprasegmentals will facilitate the development of intelligible speech. When low-frequency speech energy is used at an optimal level (dB) and accompanied by daily auditory training, Guberina contends that low-frequency speech energy does not mask high-frequency speech energy, but rather facilitates the perception of it. It is possible that speech energy in the low-frequency range may be contributing to the speech perception and production abilities of prelingually deaf children who are in the process of developing their initial speech skills. Appendix A elaborates on the points cited above.

The present study attempted to settle this controversy over the possible harmful or beneficial effects of low-frequency amplification when used for daily auditory training of young deaf children.

The Verbo-tonal method was selected for training the children because it emphasizes the development of both speech perception and speech production. In order to evaluate the progress of the children, perception and production were tested simultaneously. A similarity scale was used to assess the speech samples, so that speech production could be evaluated from the sub-intelligible to intelligible levels. It was assumed that improvements in speech production reflected changes in perception. The speech samples were also analyzed acoustically to evaluate changes in the suprasegmental aspects of the speech patterns.

METHOD

A. Equipment

Two auditory training units were used in this study: the Warren, model T-2, and the Suvag, model I. The Warren unit was equipped with one microphone (Shure, model 777); the Suvag unit had four piezoelectric microphone cartridges (Astatic, model MC-151). The latter included four microphones in parallel to improve the low-frequency response. The filters in the Suvag unit were not used during classroom activities or during the evaluation of the unit.

Each training unit independently drove five pairs of earphones (Koss, model SP-3XC) and five bone vibrators (Vibar Suvag, model 73). The bone vibrators were used in addition to the earphones because they are part of the training procedure of the Verbo-tonal method. After the first year, the Koss SP-3XC earphones were replaced by Koss K-6 earphones because the latter furnished a more comfortable fit, a better seal, and improved the low-frequency response.

The frequency responses of the two units are presented in Figure 1 where acoustic output through a Koss K-6 earphone is plotted as a function of acoustic input. The two units differed in low-frequency amplification. The Warren unit amplified from 200 to 5000 Hz, the Suvag unit amplified from 20 to 5000 Hz. Therefore, the Suvag unit provided an additional amplification of the frequencies below 200 Hz.

The Warren unit was a compression limiting system; the Suvag unit was a peak limiting device. In this study both units were used below their limiting levels operating in their linear range of input-output characteristics. The electro-acoustical parameters of both units were measured with standard measuring laboratory equipment. The description and the results of these measurements are given in Appendix B. The performance of both units was basically identical with the exception of the difference in the low-frequency region indicated above.

Two different wearable hearing aids were selected for home use for the children. The aids with the narrower frequency response (Zenith Vocalizer II) were assigned to the Warren group, and the aids with the wider frequency response (Mini Suvag) were assigned to the Suvag group. All were fitted monaurally as body type aids. Each Zenith aid used a Zenith model Y5R receiver and each Suvag aid used an Oticon model M1 receiver.

Figure 2 displays the acoustic output of the two aids as measured in a sound-treated booth with a standard test system. The difference between the two aids was similar to the difference between the two training units. The Mini Suvag hearing aid provided additional amplification in the low-frequency region below 200 Hz. Appendix B displays additional specifications of the two aids.

The Verbo-tonal audiometer was used for filtered speech testing. This included a pre-recorded tape of filtered logatomes (nonsense syllables) for obtaining detection thresholds. An Interim Report (Asp, 1972) published during the course of the project showed diagrams and provided specifications of the equipment described in this section.

BEST COPY AVAILABLE

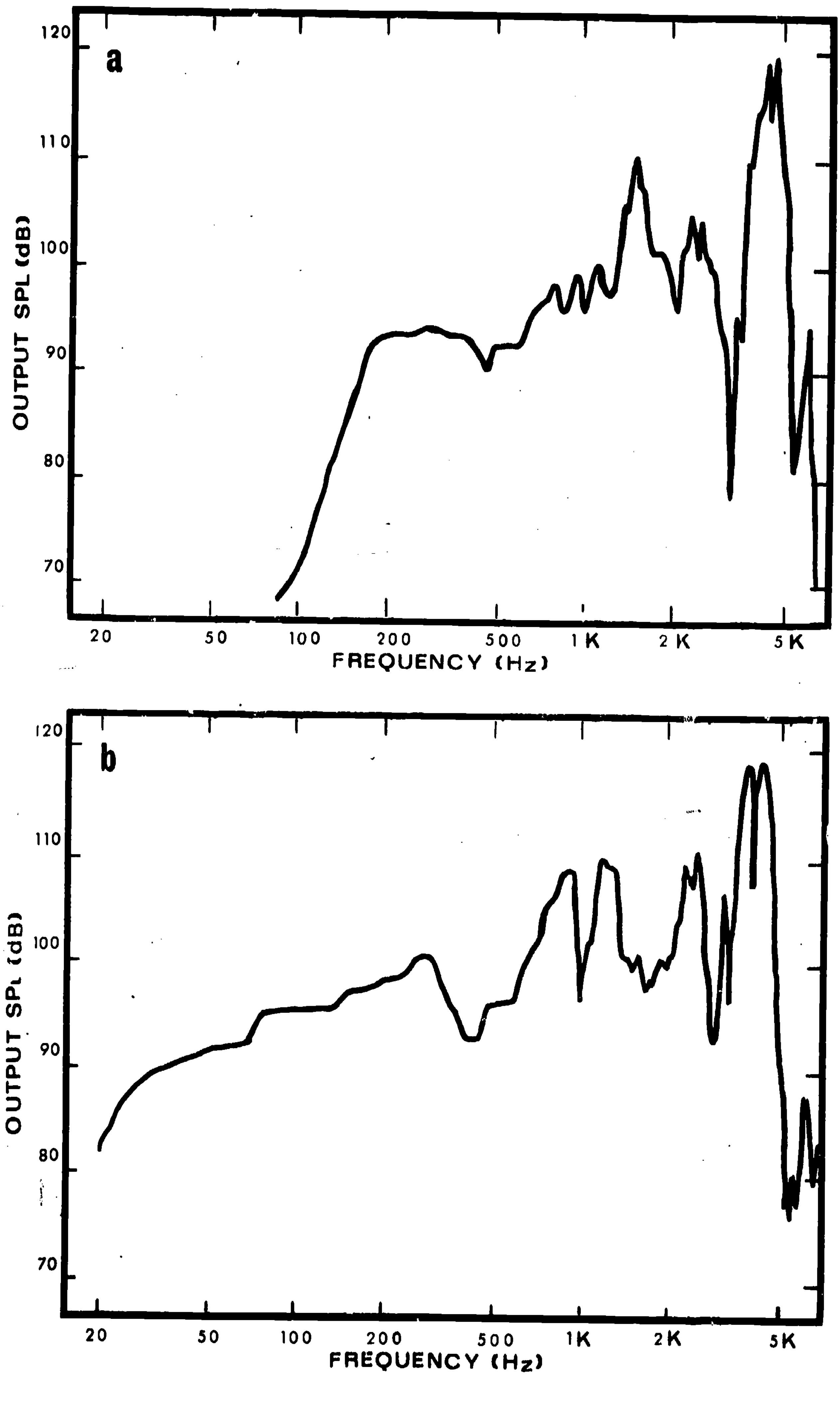


Figure 1: Frequency Responses of the Auditory Training Units Measured with 60dB Input SPL: a. Warren T-2, b. Suvag I

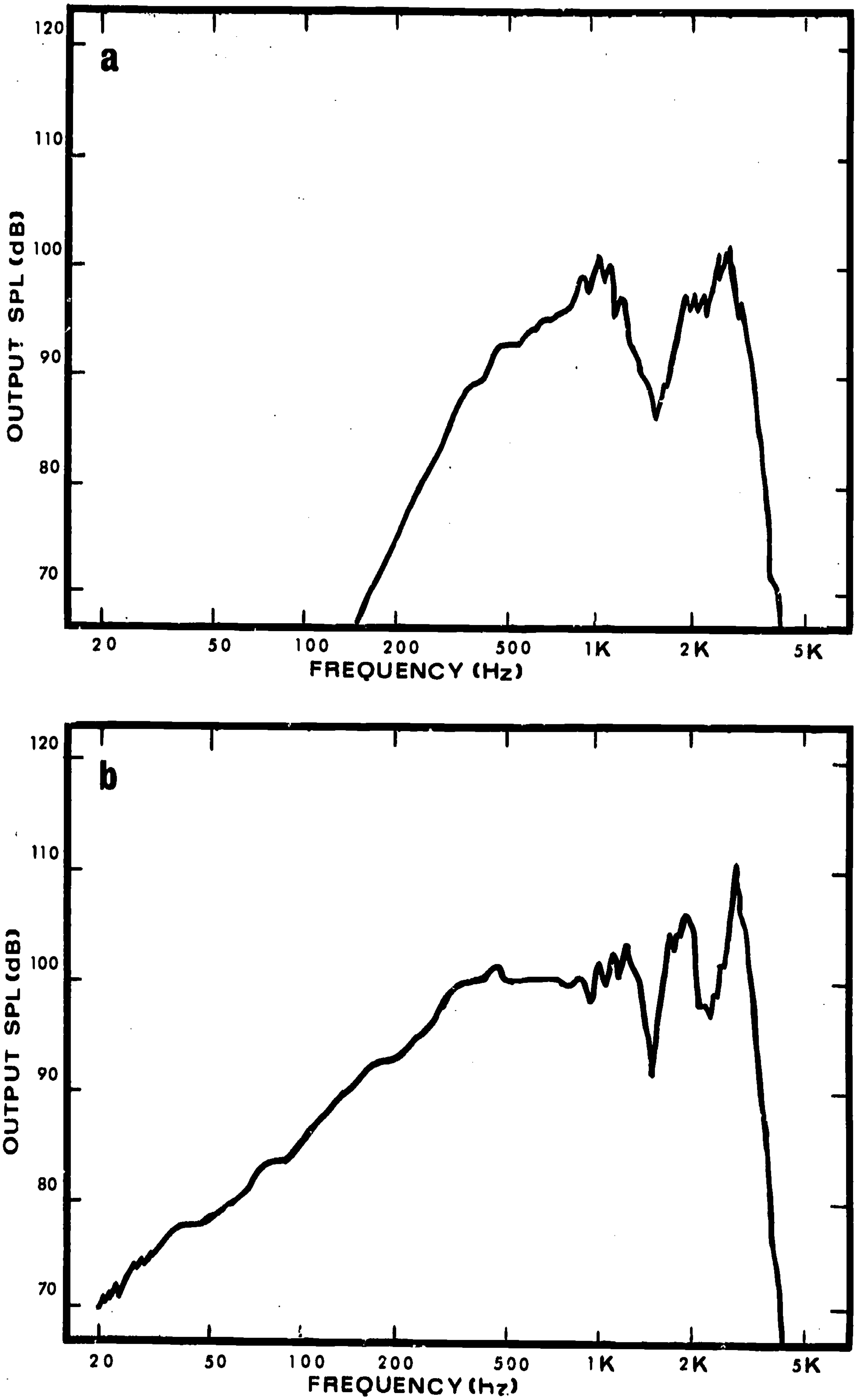


Figure 2: Frequency Responses of the Hearing Aids Measured with 60 dB Input SPL: a. Zenith Vocalizer b. Mini Suvag

B. Subjects

Throughout the three years of this study, forty-one children were evaluated and accepted into the program. The subjects were matched and assigned to either the Warren or the Suvag group. The criteria for matching included: 1) hearing-threshold level (HTL), 2) auditory perception and speech production, 3) intelligence, 4) age, 5) additional diagnostic information, and 6) sex. The first two criteria received the greater emphasis.

During the first and second years, 13 Warren subjects and 15 Suvag subjects were admitted. The period of enrollment for each subject is identified in Appendices C-1 and C-2, respectively. Sufficient data were not available to warrant inclusion of 13 additional subjects, nine of whom were admitted in the third year.

In order for a subject to be included in the final analyses, it was necessary to complete: 1) at least 20 training hours prior to the first test, and 2) at least three successive tests at which tape recorded samples were obtained. Because of the high attrition and the criteria described above, the number of subjects was reduced to 20. Thus, the match between the groups was slightly different than was initially intended.

For the final analyses, 11 Warren and 9 Suvag subjects were available. Table 1 displays diagnostic information on these subjects. The Warren subjects are identified as W1, W2, ... W11, and the Suvag subjects as S1, S2 ... S9. Within each column in Table 1, the number within the parentheses identifies the ranked position of each subject relative to the 20 subjects in this study. For each group, the means $(\overline{X})$, standard deviations (SD), and ranges are listed.

The mean 3-frequency average (0.5, 1, and 2 kHz) for the better ear was 91.6 dB for the Warren group and 88.5 dB for the Suvag group. The ranges were 53 to 110+ dB and 68 to 100 dB, respectively. The difference between the groups was 3 dB. When the 20 subjects in Table 1 were ranked from the least to the greatest hearing loss, the mean rank was 11 for the Warren group and 9 for the Suvag group. As another point of comparison, an auditory classification system (Risberg and Martony, 1970) ranging from A1 to D5 was used for the better thresholds of each subject. The mean rank of these classifications was 12 for the Warren group and 9 for the Suvag group.

The pure-tone air- and bone-conduction thresholds for the 11 Warren and 9 Suvag subjects are displayed in Tables 2-a and 2-b, respectively. The better-ear air-conduction threshold for the seven test frequencies for each subject served as the criterion measure for a one-factor analysis of variance. The results indicated that the Warren and Suvag groups were not significantly different ($F = 0.23$; $df = 1,18$, $p < 0.64$). The mean threshold of the seven test frequencies for the Warren group (85.3 dB) was 2.9 dB greater than the threshold for the Suvag group (82.4).

Table 1. Subjects used for the final analyses* BEST COPY AVAILABLE

Subject Number	Age at First Test (year-month)	Sex	3-Frequency Average (dB HTL)	Audiogram Class. Risberg-Martony	IQ	Etiology	Time of Completed Tests: 1st Year ** a-b-c	2nd Year ** a-b-c	3rd Year *** a-b	Cumulative Therapy Hours: Test 3	Test 4	Last Test
Warren:												
W1	3-7 (12.5)	M	92 (11)	B4 (7.5)	63 (20)	Premature	1-2	3-4-5	6	96(20)	238(12)	330(9)
W2	2-4 (17.5)	F	93 (12.5)	C4 (15)	133 (2)	Mat.Xray	1-2	3-4-5	6-7	189(14)	328(9)	608(6)
W3	5-4 (2)	M	80 (6)	C2 (10)	80 (17)	Heredity	1-2	3-4		374(2)	475(2)	
W4	3-11(8.5)	F	90 (9.5)	B4 (7.5)	126 (3)	Rubella	1-2-3	4-5-6	7-8	400(1)	498(1)	1276(1)
W5	4-9 (3)	F	53 (1)	A2 (1)	98 (10.5)	Unknown	1-2-3	4-5		329(4)	381(6.5)	485(7)
W6	4-4 (5.5)	F	110+(19.5)	D5 (19)	96 (12.5)	High Fever	1-2	3		190(13)		
W7	3-7 (12.5)	M	110+(19.5)	D5 (19)	87 (15)	Meningitis		1-2	3-4	211(12)	256(11)	
W8	2-4 (17.5)	M	88 (7.5)	C3 (12.5)	155 (1)	Rubella		1-2	3-4	260(7)	381(6.5)	
W9	2-2 (19)	M	77 (4)	B3 (5)	124 (4)	Meningitis		1	2-3	230(9)		
W10	4-7 (4)	M	107 (17)	C5 (16.5)	76 (18)	Unknown		1	2-3	179(15)		
W11	4-4 (5.5)	M	108+(18)	C5 (16.5)	111 (6.5)	Unknown		1	2-3	169(16)		
Mean	3-9 (10)		>92 (11)	(12)	105 (10)					239(10)	365(7)	675(6)
Std.Dev.	1-1	M=7	17		28					93	100	417
Range	2-2 to 5-4	F=4	53 to 110+	A2-D5	63 to 155					96-400	238-498	330-1276
Suvag:												
S1	3-11 (8.5)	F	78 (5)	B2 (3.5)	90 (14)	Rubella	1-2-3	4-5-6	7-8	258(8)	324(10)	1054(2)
S2	3-4 (14.5)	F	88 (7.5)	C3 (12.5)	111 (6.5)	Rubella	1-2	3-4-5	6-7	223(10)	391(5)	809(3)
S3	4-0 (7)	M	98 (14)	B4 (7.5)	98 (10.5)	Premature	1-2	3-4-5		217(11)	373(8)	464(8)
S4	5-7 (1)	M	73 (3)	B2 (3.5)	85 (16)	Rubella	1	2-3-4	5-6	337(3)	463(3)	688(5)
S5	3-10 (10)	F	105 (16)	D5 (19)	115 (5)	Rubella	1	2-3-4	5-6	292(6)	426(1)	727(4)
S6	1-8 (20)	F	90 (9.5)	B4 (7.5)	103 (9)	Meningitis		1-2-3	4-5	137(19)	153(4)	226(10)
S7	3-4 (1.5)	M	102 (15)	C3 (12.5)	64 (19)	Unknown		1-2	3-4	149(18)	235(13)	
S8	2-11 (16)	F	93 (12.5)	C3 (12.5)	96 (12.5)	High Fever		1	2-3	161(17)		
S9	3-9 (11)	F	68 (2)	A4 (2)	107 (8)	Premature		1	2-3	307(5)		
Mean	3-8 (10)		89 (9)	(9)	97 (10)					231(11)	338(6)	661(5)
Std.Dev.	1-5	M=3	13		16					73	110	286
Range	1-8 to 5-7	F=6	68 to 105	A4-D5	64 to 115					137-337	153-463	226-1054

*The number in parentheses represents the rank of each subject relative to both groups for the characteristic. The lowest rank was assigned to the condition judged to be most favorable for outcome of therapy.

** a - November, b - March, c - July

*** a - November, b - April

Table 2-a. Pure-tone air- and bone-conduction thresholds in dB HTL (ISO 1964) for the Warren group

Subject Number	Ear	Test Frequency in Hz 125	250	500	1K	2K	3K	4K	8K
W1	AC-R*	65	70	80	90	105	110+	110+	90+
	AC-L	70	85	90	95	110	110+	110+	90+
	BC-R		40	65	65+	65+		65+	
	BC-L*		35	60	65+	65+		65+	
W2	AC-R*	65	75	80	100	100	110	110	90+
	AC-L	75+	85	95	110+	100	110	105	90+
	BC-R*		35	60	65+	65+		65+	
	BC-L		35	60	65+	65+		65+	
W3	AC-R*	65	75	85	80	75		70	70
	AC-L	75+	90	85	90	85		80	85
	BC-R*		35	65	65+	65		65	
	BC-L		35	65	65+	65		65	
W4	AC-R*	65	70	75	85	110	110	110+	90+
	AC-L	60	45	75	90	105	110	110+	90+
	BC-Rx		35	60	65+	65+		65+	
	BC-L		35	60	65+	65+		65+	
W5	AC-R*	35	35	40	55	65		65	60
	AC-L	35	40	55	75	70		75	55
	BC-R*		35	45	55	65		65	
	BC-L		35	55	60	65		65	
W6	AC-R*	65	80	110	110+	110+		110+	90+
	AC-L	75	90	110	110+	110+		110+	90+
	BC-R*		35	60	65	65+		65+	
	BC-L		35	60	65+	65+		65+	
W7	AC-R*	70	85	110+	110+	110+		110+	90+
	AC-L	75+	90	110+	110+	110+		110+	90+
	BC-Rx		30	60	65+	65+		65+	
	BC-L		30	60	65+	65+		65+	
W8	AC-R	75	80	95	95	95	90	80	90
	AC-L*	75+	85	85	95	85	85	80	75
	BC-R		35	60	65+	65+		65+	
	BC-Lx		35	60	65+	65+		65+	
W9	AC-R	70	90	110	110+	110+	110+	110+	90+
	AC-L*	70	75	80	75	75	110	105	90+
	BC-Rx		35	60	65+	65+		65+	
	BC-L		35	60	65+	65+		65+	
W10	AC-R*	75+	90	100	110	110	110+	110	90+
	AC-L	75	85	110	110+	110+	110+	110+	90+
	BC-Rx		30	55	65+	65+		65+	
	BC-L		35	60	65+	65+		65+	
W11	AC-R	75	85	105	110+	110+	110+	110+	90+
	AC-L*	75	90	105	110	110+	110+	110+	90+
	BC-Rx		35	65	65+	65+		65+	
	BC-L		35	65	65+	65+		65+	
$\bar{X}$ =	AC-Better	65	72	>86	>93	>95	>106	>99	>84
	AC-Poorer	>70	82	>95	>100	>102	>107	>101	>87
	BC-Better		34	59	>64	>65		>65	
	BC-Poorer		35	61	>65	>65		>65	

KEY: AC = Air Conduction
BC = Bone Conduction
R = Right
L = Left

* = Better Ear at .5, 1, and 2 kHz
x = Both Ears the Same at .5, 1, and 2 kHz
> = Greater Than
+ = No Response at Limit of Audiometer

BEST COPY AVAILABLE

Table 2-b. Pure-tone air- and bone-conduction thresholds in dB HTL (ISO 1964) for the *Suvag* group

Subject Number	Ear	Test Frequency in Hz							
		125	250	500	1K	2K	3K	4K	8K
S1	AC-R	75	85	105	110	100	100	105	90+
	AC-L*	55	60	80	85	70	75	70	55
	BC-Rx		30	60	65+	65+		65+	
	BC-L		30	60	65+	65+		65+	
S2	AC-R*	75+	90+	95	85	85		75	80
	AC-L	75+	90+	100	105	110+		110+	90+
	BC-Rx		30	55	65+	65+		65+	
	BC-L		30	55	65+	65+		65+	
S3	AC-R*	75+	75	95	95	105		110+	90+
	AC-L	60	50	95	100	105		110+	90+
	BC-Rx		35	60	65+	65+		65+	
	BC-L		35	60	65+	65+		65+	
S4	AC-R	65	75	90	100	110+	110+	110+	90+
	AC-L*	65	70	80	75	65	60	55	50
	BC-R		35	60	65+	65+		65+	
	BC-L*		30	55	65	55		35	
S5	AC-R*	75	85	105	110	100	110+	110+	90+
	AC-L	75+	90+	110	110+	110+	110+	110+	90+
	BC-R*		35	50	65+	65+		65+	
	BC-L		35	60	65+	65+		65+	
S6	AC-R	65	80	90	90	105		110+	90+
	AC-L*	60	65	75	90	105		105	90+
	BC-Rx		35	60	65+	65+		65+	
	BC-L		35	60	65+	65+		65+	
S7	AC-R*	70	85	105	105	95	105	105	90+
	AC-L	75	85	105	110	100	95	95	90+
	BC-R		35	60	65+	65+		65+	
	BC-L*		30	55	65+	65+		65+	
S8	AC-R	75	90	110	95	90	95	110	90
	AC-L*	75+	85	95	95	90	100	110+	90+
	BC-R*		35	55	65+	65+		65+	
	BC-L		35	60	65+	65+		65+	
S9	AC-R*	35	40	45	55	105	110	110	90+
	AC-L	40	50	50	55	105	110	105	90+
	BC-Rx		35	45	55	65+		65+	
	BC-L		35	45	55	65+		65+	
$\bar{X}$ = AC-Better		>63	>70	86	88	91	>91	>93	>81
AC-Poorer		>69	>80	95	>97	>104	>106	>109	>90
BC-Better			33	55	>64	>64		>62	
BC-Poorer			34	58	>64	>65		>65	

KEY: AC = Air Conduction
BC = Bone Conduction
R = Right
L = Left

* = Better Ear at .5, 1, and 2 kHz
x = Both Ears the same at .5, 1, and 2 kHz
> = Greater Than
\+ = No Response at Limit of Audiometer

As measured by the Leiter International Performance Scale, the mean IQ was 104.5 for the Warren group and 96.6 for the Suvag group. The ranges were 63 to 155 and 64 to 115, respectively. The mean ranks were the same for both groups. Table 1 displays the measures.

At the first test of each subject, the mean ages were 3 years, 9 months, for the Warren group, and 3 years, 8 months, for the Suvag group with ranges of 2 years, 2 months to 5 years, 4 months, and 1 year, 8 months to 5 years, 7 months, respectively. There were 7 males and 4 females in the Warren group; the Suvag group had 3 males and 6 females. The etiologies of the two groups varied slightly.

During the three years of the project, the time period for completing successive tests was slightly different for each subject. Table 1 identifies these time periods and the corresponding cumulative training hours. There were three tests in both the first and second years in November, March and July, and two tests in November and April in the third year. The first test of the first year was given after two months of training. Each of the remaining tests (Test 2 to Test 8) were given after four-month periods of training. The test sequence unique to each subject was identified by the test numbers. These numbers indicate the number of consecutive four-month training periods which ended with the tests. For example, for W4, the first test was Test 1 and the last was Test 8.

The number of cumulative training hours was different for each subject. The group means at Test 3 were 239 hours for the Warren group, and 231 hours for the Suvag group. The ranges were 69 to 400, and 137 to 337 hours, respectively.

Other characteristics of this population were: 1) 60% of the children were from low-income families, and 2) 90% were transported daily from distances between 25 and 100 miles of the Center. Appendix D displays information on the subjects, based on judgements of the classroom teachers.

In summary, the subjects comprising the Warren group were slightly older, had a higher IQ, and had more training hours, whereas the Suvag group had 3 dB more sensitivity to pure tones.

C. Teachers and the Tester

Nine teachers were employed for various lengths of time during the three-year project. The median age was 25 years, with a range of 22 to 50 years. Eight female teachers had a mean fundamental frequency (F_0) of 202 Hz with a range of 190 to 210 Hz. The one male teacher had an F_0 of 110 Hz. Three teachers had a General American dialect and six had a Southern dialect. Appendix E provides the following additional information about the teachers: a) college degree(s), b) major, c) prior teaching experience, d) state(s) where raised, and 3) percentage of employment.

The supervisor, who was also the tester, trained the eight teachers to use the principles of the Verbo-tonal method (see Appendix F) in the classroom, so that the therapy procedures would be similar for each class. Appendix G identifies the nature of training and the supervision of the classroom teachers. A comprehensive description of this method can be obtained from the video-tapes and the articles listed in Appendix H.

The tester had a mean F_0 of 200 Hz for conversational speech, and a mean F_0 of 252 Hz when presenting the test words. Other characteristics of her voice are given in a later section describing the acoustic measures.

D. Experimental Design

Each child was trained in a classroom situation through the appropriate training unit, required to wear the assigned hearing aid at home, and tested after a four-month period of training.

Three classrooms were instrumented with the Warren and the Suvag training units that were described in the previous Equipment section. A block diagram of the classroom instrumentation is displayed in Figure 3. The microphones of the two units were secured in one fixture, so a single teacher could speak simultaneously through two different training units. The Warren group received a narrower frequency response than the Suvag group. Both units were concealed in closets within the classroom. The classrooms were approximately 19' X 12' with hardwood floors, plastered walls and ceilings, and curtains on all windows.

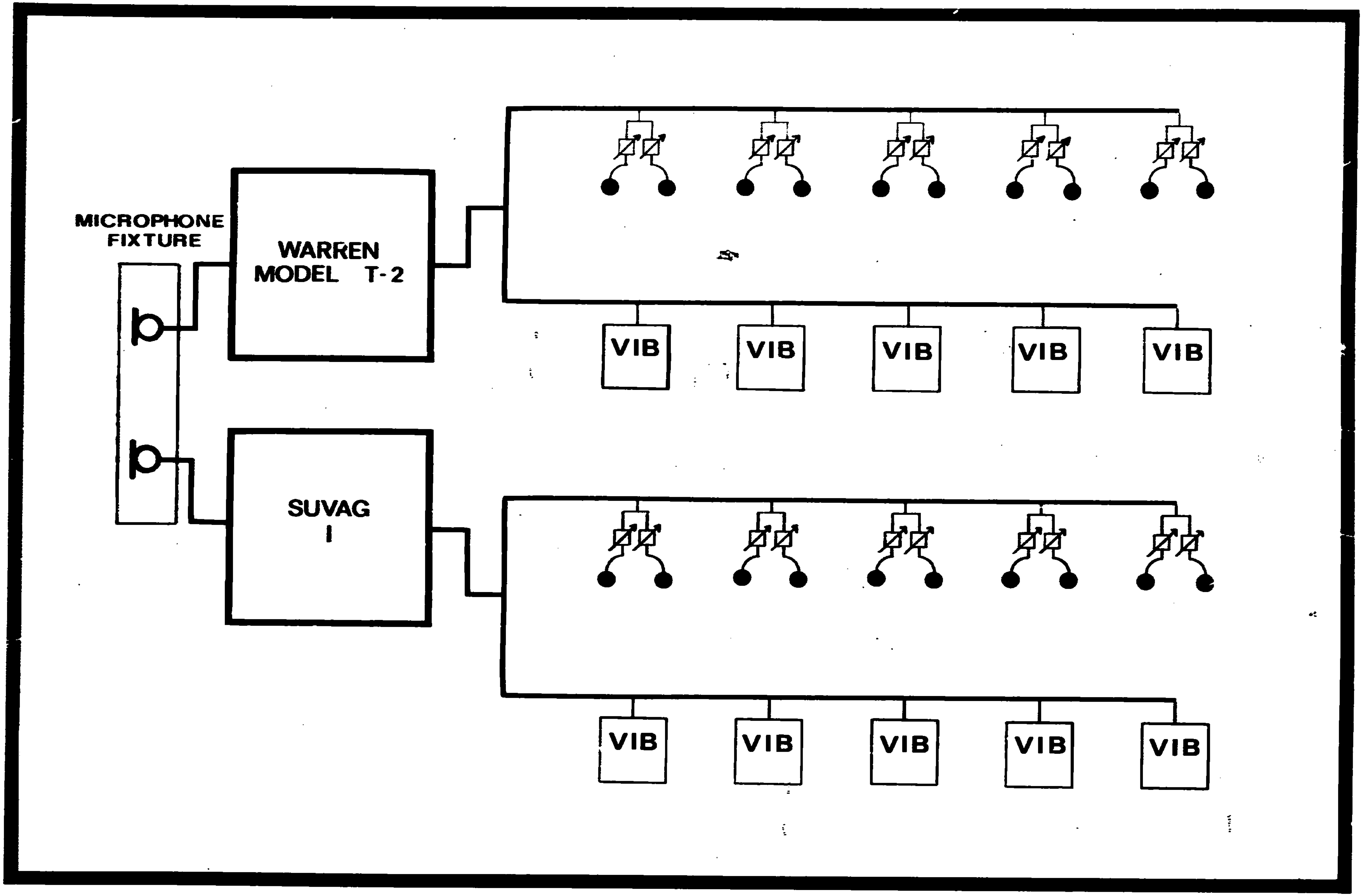
Within each classroom, each earphone and each vibrator had an individual control setting (attenuator) to control the output level (see Figure 3). Each training unit was calibrated with a 1000 Hz input tone at 82 dB SPL (re 2×10^{-4} dyne/cm^2). The amplifiers of the Warren and Suvag units were adjusted and fixed to obtain an acoustic output of 122 dB SPL at the earphones with the earphone attenuators at the maximum position corresponding to the minimum attenuation. Table 3 displays the gain and the output level in dB as a function of the control setting as measured for the calibration tone.

Table 3. Gain and output level in dB as a function of the control setting.

Control Setting	Min	1	2	3	4	5	6	7	8	9	10	Max
Gain dB	-10	14	18	22	24	26	29	32	35	37	39	40
Output SPL (dB)	72	96	100	104	106	108	111	114	117	119	121	122

The control settings for the vibrator were eliminated during the third year of the program to secure the maximum output level for the Suvag vibrators. These vibrators were capable of producing higher output levels than conventional vibrators. They had a low inertia and a good response at low frequencies.

Figure 3: Classroom Instrumentation



Within each classroom the training was done in daily group sessions in accordance with the Verbo-tonal procedures. These procedures consisted of body movements and implementations, rhythmical stimulation, reading readiness, and individual work. A detailed description of the procedures is presented in Appendix F.

The number of children within each of the three classrooms ranged between three and six, with an attempt to have an equal number from the Warren and Suvag groups. The children were re-evaluated periodically to determine if reassignment to another class would provide optimal learning. Changes in class assignment usually occurred three times each year. All children attended group sessions. The older children were scheduled for three hours daily, or 15 hours per week. The younger ones were seen 1-1/2 hours daily, or seven hours per week. The number of hours for a class of younger children was increased throughout the year as the group became capable of profitting from the stimulation.

The children wore binaural earphones for approximately 75% of the training time. There was an attempt to work at the most comfortable listening level for each ear of each child. However, at the outset of training, some of the children were unsophisticated listeners and could not select proper listening levels. Therefore, the levels were set by the audiologist and the classroom teacher based on the child's audiogram and the child's speech perception in the classroom. The classroom teacher checked the individual setting daily to insure that the settings were neither too low nor too high for the child. When the child had developed the sophistication to select the most comfortable listening level, he was allowed to make the adjustments of the controls, if the settings did not differ greatly from the estimated levels and if a child was consistent in the selection. As the child improved in perception, he tended to use a lower setting. The classroom teacher checked each training unit daily to insure optimum performance.

After the study was completed, each teacher was requested to select the settings of each ear of each child from the daily records. The mean setting for the Warren group was 116 dB SPL (for 82 dB input), and it was 114 dB SPL for the Suvag group. The ranges were 103 dB to 109 dB, and 105 dB to 121 dB, respectively.

The classroom teacher was responsible for teaching the child and keeping the parents informed as to how to assist the child in learning. The latter information did not include formal therapy techniques. Social workers were available for both group and individual counseling. The parents were not informed about the nature of the study.

Each child was fitted with a Zenith or a Mini Suvag aid for home use. This usually occurred within the first three months of therapy. The children were encouraged to wear their aids after classroom activities had terminated for the day.

The assumption of this training program was that a child's progress in speech perception will be reflected in his progress in speech production. Therefore, the evaluation at the end of each four-month training period used the speech samples that were produced by the child.

Twenty-seven one- and two-syllable words were selected that were common to preschool children and had phonemes that corresponded to the principle of phonetic progression. The concept of phonetic progression is described in the Verbo-tonal procedures in Appendix F. Fifteen of these words (a modified list) that were representative of the larger list were used initially with the younger children who could not attend to the longer list. Preliminary statistical analysis revealed similar results using either 15 or 27 words. Thus, the 15-word list was used for all statistical analyses, for it was common to all children. The 15-word list is presented in Table 4.

Table 4. Fifteen one- and two-syllable words used for testing the subjects

1. puppy	6. bee	11. cookie
2. pillow	7. mama	12. cheek
3. bunny	8. toe	13. lamb
4. baby	9. daddy	14. shoe
5. bye bye	10. come	15. soup

Speech samples were obtained at four-month intervals, by tape recording both the tester's stimulus and the child's response during the test sessions. The same tester who had used this testing procedure for two years prior to the project tested each child. The tester presented each word twice in succession and a sufficient time was allowed for response following each presentation. The second presentation of each word was to allow the child to respond if the first presentation was not perceived.

The test words were presented under four testing conditions:

1) without visual clues and without amplification ($\overline{VA}$)
2) without visual clues and with amplification ($\overline{V}A$)
3) with visual clues and without amplification ($V\overline{A}$)
4) with visual clues and with amplification (VA)

The words were presented randomly for each test condition. The children were familiar with the conditions of with and without (unaided training) amplification for they were part of the daily classroom situation. Appendix I describes the testing procedure in greater detail.

RESULTS

A. Similarity Scale

The speech production of the children was often unintelligible and could not be evaluated by standard testing procedures for evaluating speech intelligibility. Thus, a new procedure was developed and it was identified as the similarity scale.

Since each test word was presented twice by the tester, the "better" of two responses of each child was selected as the representative speech sample for each test word. These samples (stimulus and response) were randomized, re-recorded, and judged on a 9-point similarity scale by a panel of 12 to 15 normal-hearing college students.

The listeners were instructed to judge each child's response as it related to the teacher's stimulus. Number 1 on the scale indicated that the response was similar to the stimulus, and number 9 indicated that it was dissimilar. Numbers between 1 and 9 represented degrees of similarity. After a practice session, the listeners responded by circling a number that represented their judgment for each speech sample.

As an estimate of intra-judge reliability, 15 listeners judged 20 speech samples twice, at the outset and at the end of a listening session. For these listeners, the mean correlation coefficient was 0.88 with a range of 0.77 to 0.98. This suggested a high degree of intra-judge reliability.

The mean rating of the panel of listeners was computed for each test word. If a child did not respond to a test word, a value of 9 was assigned to that word; thus, each word was represented by a similarity score. For each subject, the criterion measure was the mean similarity score of the 15-word list for each test condition.

The mean similarity scores for 11 Warren (W) and 9 Suvag (S) subjects are displayed in Tables 5-a, 5-b, and 5-c for the consecutive tests. After Test 3, the number of subjects decreased as tests increased. As a result, the statistical analysis to compare the groups will be confined to the first three tests (see Table 5-a) where 11 Warren and 9 Suvag subjects had similarity score for the four test conditions.

A four-factor analysis of variance (Winer, 1971) with repeated measures on three factors (tests, amplification, and visual clues) was used to analyze the criterion measure. There was no statistically significant difference between the similarity score for the Suvag (6.9) and the Warren (7.4) groups ($F = 0.39$; $df = 1,18$; $p < 0.55$). There was a significant improvement (a decrease in similarity scores) ($F = 13.75$; $df = 2,36$; $p < 0.0001$) for both groups over the three tests (Test 1 = 7.6, Test 2 = 7.1, and Test 3 = 6.8). The testing conditions <u>with</u> amplification (7.1) were slightly better than <u>without</u> amplification (7.2), but the difference was not statistically significant ($F = 3.67$; $df = 1,18$; $p < 0.07$). The testing condition <u>with</u> visual clues (7.0) was significantly better ($F = 13.27$; $df = 1,18$; $p < 0.002$) than <u>without</u> visual clues (7.3). The interactions were not significant.

The testing condition of amplification <u>without</u> visual clues ($\overline{V}A$) was considered the most sensitive test for the effects of low-frequency amplification. A separate analysis of variance was computed for this condition, excluding subject W5. This subject was a hard-of-hearing child (53 dB HTL) who had an initial similarity score equal to 1.6, the lowest of any subject. The Suvag group (6.8) had a better score than the Warren group (8.0); however, it was not statistically significant ($F = 2.77$; $df = 1,17$; $p < 0.09$).

Table 5-a. Mean similarity score of 15 words for four conditions ($\overline{V}\overline{A}$, $\overline{V}A$, $V\overline{A}$, VA) as functions of Warren and Suvag Groups and Tests 1-3*.

BEST COPY AVAILABLE

Subject No.	Test 1 $\overline{V}\overline{A}$	Test 1 $\overline{V}A$	Test 1 $V\overline{A}$	Test 1 VA	Test 2 $\overline{V}\overline{A}$	Test 2 $\overline{V}A$	Test 2 $V\overline{A}$	Test 2 VA	Test 3 $\overline{V}\overline{A}$	Test 3 $\overline{V}A$	Test 3 $V\overline{A}$	Test 3 VA
Warren Group:												
W1	9.0(17)	9.0(16.5)	9.0(17.5)	9.0(17)	8.2(12)	8.4(12)	9.0(18)	8.4(15)	8.6(17)	8.7(17)	9.0(20)	8.6(17)
W2	8.7(12)	9.0(16.5)	8.9(13.5)	9.0(17)	8.4(14.5)	8.5(13.5)	8.3(14)	7.7(10)	8.2(14)	8.4(15)	8.5(16)	7.7(12)
W3	7.1(4)	6.4(4)	5.8(4)	5.7(4)	6.5(5)	5.6(4)	5.9(4)	5.5(5)	6.9(6)	6.9(7)	6.2(7.5)	5.0(5)
W4	9.0(17)	9.0(16.5)	9.0(17.5)	9.0(17)	8.2(12)	7.9(8)	6.7(6)	7.7(10)	7.3(8)	6.0(5)	6.1(6)	6.3([illegible])
W5	1.5(1)	1.6(1)	1.[illegible](1)	1.7(1)	1.7(1)	1.6(1)	1.3(1)	1.4(1)	1.6(1)	1.4(1)	1.7(1)	1.4(1)
W6	7.8(5.5)	8.0(7.5)	7.8(10)	7.5(7)	8.4(14.5)	8.6(15.5)	7.7(13)	7.9(13)	7.5(9)	7.6(11)	7.2(11)	7.8(13.5)
W7	8.4(10)	8.1(9)	6.9(5.5)	7.9(10.5)	7.9(8.5)	8.0(9.5)	7.0(8)	6.6(6)	8.0(13)	7.5(10)	6.9(9.5)	6.8(9)
W8	9.0(17)	9.0(16.5)	9.0(17.5)	9.0(17)	9.0(19)	9.0(19)	9.0(18)	9.0(19)	8.5(16)	8.4(15)	8.9(18.5)	7.9(15)
W9	9.0(17)	9.0(16.5)	9.0(17.5)	9.0(17)	9.0(19)	9.0(19)	9.0(18)	9.0(19)	8.9(18.5)	9.0(19)	8.3(14)	9.0(19)
W10	8.6(11)	8.5(11.5)	8.0(11)	8.2(12)	7.9(8.5)	8.5(13.5)	7.5(10.5)	8.2(14)	7.8(11)	7.9(13)	7.4(12)	8.4(16)
W11	9.0(17)	8.5(11.5)	8.7(12)	7.7(8)	8.2(12)	8.2(11)	7.5(10.5)	7.8(12)	7.8(11)	8.4(15)	7.7(13)	7.8(13.5)
$\overline{X}$ =	7.9(11.7)	7.8(11.5)	7.6(11.5)	7.6(11.6)	7.6(11.5)	7.6(11.4)	7.2(11.0)	7.2(11.3)	7.4(11.3)	7.3(11.6)	7.1(11.7)	7.0(11.5)
Suvag Group:												
S1	8.1(8)	7.9(5.5)	7.6(8.5)	7.8(9)	5.6(4)	5.9(5)	6.4(5)	4.0(4)	5.2(4)	3.5(4)	5.3(4)	4.0(4)
S2	7.9(7)	8.2(10)	6.9(5.5)	6.7(5)	8.1(10)	7.3(6.5)	6.[illegible](7)	7.2(8)	7.2(7)	6.5(6)	6.0(5)	6.4(7)
S3	7.8(5.5)	7.9(5.5)	7.5(7)	7.3(6)	7.8(7)	8.0(9.5)	7.3(9)	7.7(10)	6.6(5)	7.1(8)	6.9(9.5)	7.2(10)
S4	5.1(3)	4.1(2)	4.9(2)	3.9(2.5)	4.6(3)	4.4(3)	4.3(3)	3.6(3)	4.5(3)	3.1(2)	4.0(3)	3.3(3)
S5	8.3(9)	8.0(7.5)	7.6(8.5)	7.9(10.5)	7.7(6)	7.3(6.5)	7.6(12)	6.7(7)	8.3(15)	7.8(12)	6.2(7.5)	6.5(8)
S6	8.9(13)	9.0(16.5)	8.9(13.5)	8.7(13)	8.8(17)	8.8(17)	8.8(15)	8.9(17)	8.9(18.5)	9.0(19)	8.9(18.5)	9.0(19)
S7	9.0(17)	9.0(16.5)	9.0(17.5)	9.0(17)	9.0(19)	8.6(15.5)	9.0(18)	8.8(16)	9.0(20)	9.0(19)	8.6(17)	9.0(19)
S8	9.0(17)	9.0(16.5)	9.0(17.5)	9.0(17)	8.7(16)	9.0(19)	9.0(18)	9.0(19)	7.8(11)	7.3(9)	8.4(15)	7.5(11)
S9	3.8(2)	4.3(3)	5.2(3)	3.9(2.5)	2.6(2)	3.2(2)	2.6(2)	2.2(2)	2.7(2)	3.2(3)	2.5(2)	2.2(2)
$\overline{X}$ =	7.5(9)	7.5(9.2)	7.4(9.2)	7.1(9.2)	7.0(9.3)	6.9(9.3)	6.9(9.9)	6.5(9.6)	6.7(9.5)	6.3(9.1)	6.3(9.1)	6.1(9.2)

V - visual clues
A - amplification
$\overline{V}$ - without visual clues
$\overline{A}$ - without amplification

*The value within parentheses represents the rank of each subject relative to the total number of subjects for both groups for the particular column.

Table 5-b. Mean similarity score of 15 words for four conditions ($\overline{VA}$, $\bar{V}A$, $V\bar{A}$, VA) as functions of Warren and Suvag groups and Tests 4-6*

BEST COPY AVAILABLE

Subject No.	Test 4				Test 5				Test 6			
	$\overline{VA}$	$\bar{V}A$	$V\bar{A}$	VA	$\overline{VA}$	$\bar{V}A$	$V\bar{A}$	VA	$\overline{VA}$	$\bar{V}A$	$V\bar{A}$	VA
Warren Group:												
W1	8.7(12)	8.5(12)	8.5(12)	8.4(12)	9.0(9.5)	7.9(9)	8.6(10)	8.6(10)	8.5(7)	8.5(7)	8.6(7)	8.0(7)
W2	7.9(9)	8.4(11)	8.2(11)	8.2(11)	7.6(7)	8.1(10)	8.0(9)	7.0(7)	8.4(6)	7.8(6)	7.2(6)	6.5(6)
W3	7.7(8)	7.0(7)	5.4(4)	5.6(5)								
W4	5.0(3)	5.3(4)	5.3(3)	4.6(4)	5.8(4)	5.3(4)	6.2(5)	3.9(4)	6.8(4)	4.9(4)	5.1(2)	4.0(4)
W5	1.1(1)	1.2(1)	1.5(1)	1.2(1)	1.2(1)	1.1(1)	1.1(1)	1.3(1)				
W6												
W7	8.1(10)	7.7(10)	6.9(8)	6.6(10)								
W8	8.2(11)	6.1(5)	7.6(10)	6.5(8.5)								
W9												
W10												
W11												
$\bar{X}$ =	6.7(7.7)	6.3(7.1)	6.2(7.0)	5.9(7.4)	5.9(5.4)	5.6(6.0)	6.0(6.3)	5.2(5.5)	7.9(5.7)	7.1(5.7)	7.0(5.0)	6.2(5.7)
Suvag Group:												
S1	4.5(2)	2.8(2)	5.8(5)	3.2(3)	4.8(2)	2.7(2)	5.3(3)	2.6(3)	3.9(2)	3.0(2)	5.6(4.5)	2.1(1)
S2	7.4(6)	7.3(8)	6.5(7)	6.5(8.5)	7.7(8)	6.7(5)	6.5(6)	7.3(8)	4.0(3)	3.7(3)	5.3(3)	3.5(3)
S3	7.4(6)	6.6(6)	7.1(9)	6.4(7)	7.5(6)	7.1(7)	6.6(7)	6.5(6)				
S4	6.3(4)	3.6(3)	4.7(2)	2.7(2)	5.4(3)	2.8(3)	2.8(2)	2.1(2)	2.9(1)	2.6(1)	2.3(1)	2.5(2)
S5	7.4(6)	7.5(9)	6.4(6)	6.3(6)	7.2(5)	6.8(6)	5.6(4)	5.9(5)	7.6(5)	6.6(5)	5.6(4.5)	5.1(5)
S6	9.0(14)	8.7(13.5)	9.0(13.5)	9.0(14)	9.0(9.5)	7.3(8)	7.6(8)	7.6(9)				
S7	8.8(13)	8.7(13.5)	9.0(13.5)	8.9(13)								
S8												
S9												
$\bar{X}$ =	7.3(7.3)	6.5(7.9)	6.9(8.0)	6.1(7.6)	6.9(5.6)	5.6(5.2)	5.7(5.0)	5.3(5.5)	4.6(2.8)	4.0(2.8)	4.7(3.3)	3.3(2.8)

*The value within parentheses represents the rank of each subject relative to the total number of subjects for both groups for the particular column.

V - visual clues
A - amplification
$\bar{V}$ - without visual clues
$\bar{A}$ - without amplification

Table 5-c. Mean similarity score of 15 words for four conditions ($\overline{VA}$, $\bar{V}A$, $V\bar{A}$, VA) as functions of Warren and Suvag groups and <u>Tests 6 and 7</u>*.

BEST COPY AVAILABLE

Subject No.	Test 7				Test 8			
	$\overline{VA}$	$\bar{V}A$	$V\bar{A}$	VA	$\overline{VA}$	$\bar{V}A$	$V\bar{A}$	VA
Warren Group:								
W1								
W2	9.0(4)	7.0(4)	5.9(4)	5.6(4)				
W3								
W4	6.1(3)	6.6(3)	5.0(2)	5.0(3)	4.5(2)	3.6(2)	4.0(2)	2.4(2)
W5								
W6								
W7								
W8								
W9								
W10								
W11								
$\bar{X}$ =	7.6(3.5)	6.8(3.5)	5.5(3)	5.3(3.5)				
Suvag Group:								
S1	4.7(1)	2.4(1)	3.0(1)	2.8(1)	2.5(1)	2.3(1)	3.3(1)	2.3(1)
S2	5.3(2)	5.2(2)	5.4(3)	4.8(2)				
S3								
S4								
S5								
S6								
S7								
S8								
S9								
$\bar{X}$ =	5.0(1.5)	3.8(1.5)	4.2(2.0)	3.8(1.5)				

*The value within parentheses represents the rank of each subject relative to the total number of subjects for both groups for the particular column.

V - visual clues
A - amplification
$\bar{V}$ - without visual clues
$\bar{A}$ - without amplification

To evaluate the progress of each subject over the training periods, the similarity scores computed for Test 3 were subtracted from the scores obtained from Test 1. The difference scores, expressing the amount of improvement on the similarity scale, are shown in Table 6. The values in parentheses are the ranks of each subject relative to the 20 subjects in both groups. Most of the difference scores were positive, indicating an improvement. The mean difference scores were better for the Suvag group for all test conditions. The ranks were better for the Suvag than for the Warren group for all four testing conditions.

As can be seen from Tables 5-b and 5-c, some subjects demonstrated noticeable improvement as the test number increased. Figures 4-a and 4-b show subjects W4 and S1, who improved more than 5 points in similarity score from Test 1 to Test 8. After Test 3, W4 scored noticeably better when both visual clues and amplification (VA) were available. With the removal of either visual clues ($\overline{V}$A) or amplification (V$\overline{A}$), W4 did poorer. On the other hand, amplification (A) was more important for S1, and after Test 2, the scores were similar with (VA) or without ($\overline{V}$A) visual clues. Both subjects improved on the most difficult testing conditions when both amplification and visual clues were not available ($\overline{VA}$). This latter observation is noteworthy since W4 and S1 had hearing losses of 90 and 78 dB HTL, respectively.

B. Acoustic Measures

The similarity scale proved to be a good measure of the improvement in speech production from the sub-intelligible to the intelligible levels. Data relative to this observation will be presented in a later section on Intelligibility. To estimate the contribution of the suprasegmental features to the correctness of the produced word, narrow- and wide-band spectrograms (Kay Elemetrics, model 6061A) were made for the speech samples (stimulus and response) for the testing condition of amplification without visual clues ($\overline{V}$A). For practical reasons it was necessary to reduce the amount of data for this type of analysis. This condition ($\overline{V}$A) was selected because it appeared to be the most sensitive measure of the frequency response of the training units.

From the measurements that were obtained from spectrograms, the following will be discussed: 1) mean fundamental frequency (F_o) in Hz; 2) intonational contour in percent correct; 3) syllable match in percent correct; 4) voice duration in msec (a later analysis at the word level will use the response/stimulus ratio); and 5) the latency in msec, for the time between the termination of the teacher's stimulus and the onset of the child's response. Three additional measures will not be included in this discussion because they appeared to be redundant with those mentioned above. These measures were: range of F_o in semi-tones, envelope match, and total duration. Most acoustic measures were taken from the narrow-band display, with the wide-band display used to confirm these judgements. A description of the measurement procedures for each parameter is provided in Appendix J. The choice of the parameters was based on studies by Lehiste (1970) and Eguchi and Hirsh (1969), and the experience of the investigators of this project.

Table 6. Difference scores for each subject on the similarity scale obtained by subtracting Test 3 from Test 1.

Subject Number	Testing Conditions				Mean
	$\overline{V}\overline{A}$	$\overline{V}A$	$V\overline{A}$	VA	
Warren Group:					
W1	0.4(12.5)	0.3(13)	0.0(18)	0.4(11)	0.28(14)
W2	0.5(10.5)	0.6(9.5)	0.4(13.5)	1.3(6)	0.70(8)
W3	0.2(15)	-0.5(20)	-0.4(20)	0.7(9)	0.00(19)
W4	1.7(2)	3.0(2)	2.9(1)	2.7(2)	2.58(2)
W5	-0.1(20)	0.2(14.5)	0.2(15)	0.3(12.5)	0.15(17)
W6	0.3(14)	0.4(12)	0.6(10.5)	-0.3(19.5)	0.25(15)
W7	0.4(12.5)	0.6(9.5)	0.0(18)	1.1(7.5)	0.53(12)
W8	0.5(10.5)	0.6(9.5)	0.1(16)	1.1(7.5)	0.58(10)
W9	0.1(16)	0.0(18)	0.7(8)	0.0(15.5)	0.20(16)
W10	0.8(7)	0.6(9.5)	0.6(10.5)	-0.2(18.5)	0.45(13)
W11	1.2(4)	0.1(16)	1.0(5)	-0.1(17)	0.55(11)
Mean	0.5(11.3)	0.5(12.1)	0.6(12.3)	0.6(11.5)	0.55(12.5)
Suvag Group:					
S1	2.9(1)	4.4(1)	2.3(3)	3.8(1)	3.35(1)
S2	0.7(8)	1.7(3.5)	0.9(6.5)	0.3(12.5)	0.90(5)
S3	1.2(4)	0.8(7)	0.6(10.5)	0.[illegible](14)	0.68(9)
S4	0.6(9)	1.0(6)	0.9(6.5)	0.6(10)	0.78(6)
S5	0.0(18)	0.2(14.5)	1.4(4)	1.4(5)	0.75(7)
S6	0.0(18)	0.0(18)	0.0(18)	-0.3(19.5)	-0.08(20)
S7	0.0(18)	0.0(18)	0.4(13.5)	0.0(15.5)	0.10(18)
S8	1.2(4)	1.7(3.5)	0.6(10.5)	1.5(4)	1.25(4)
S9	1.1(6)	1.1(5)	2.7(2)	1.7(3)	1.65(3)
Mean	0.9(9.6)	1.2(8.5)	1.1(8.3)	1.0(9.4)	1.05(8.1)

KEY: V = Visual Clues

A = Amplification

$\overline{V}$ = Without Visual Clues

$\overline{A}$ = Without Amplification

The numbers in parentheses indicate the rank of the subject's difference score relative to both groups for the test condition.

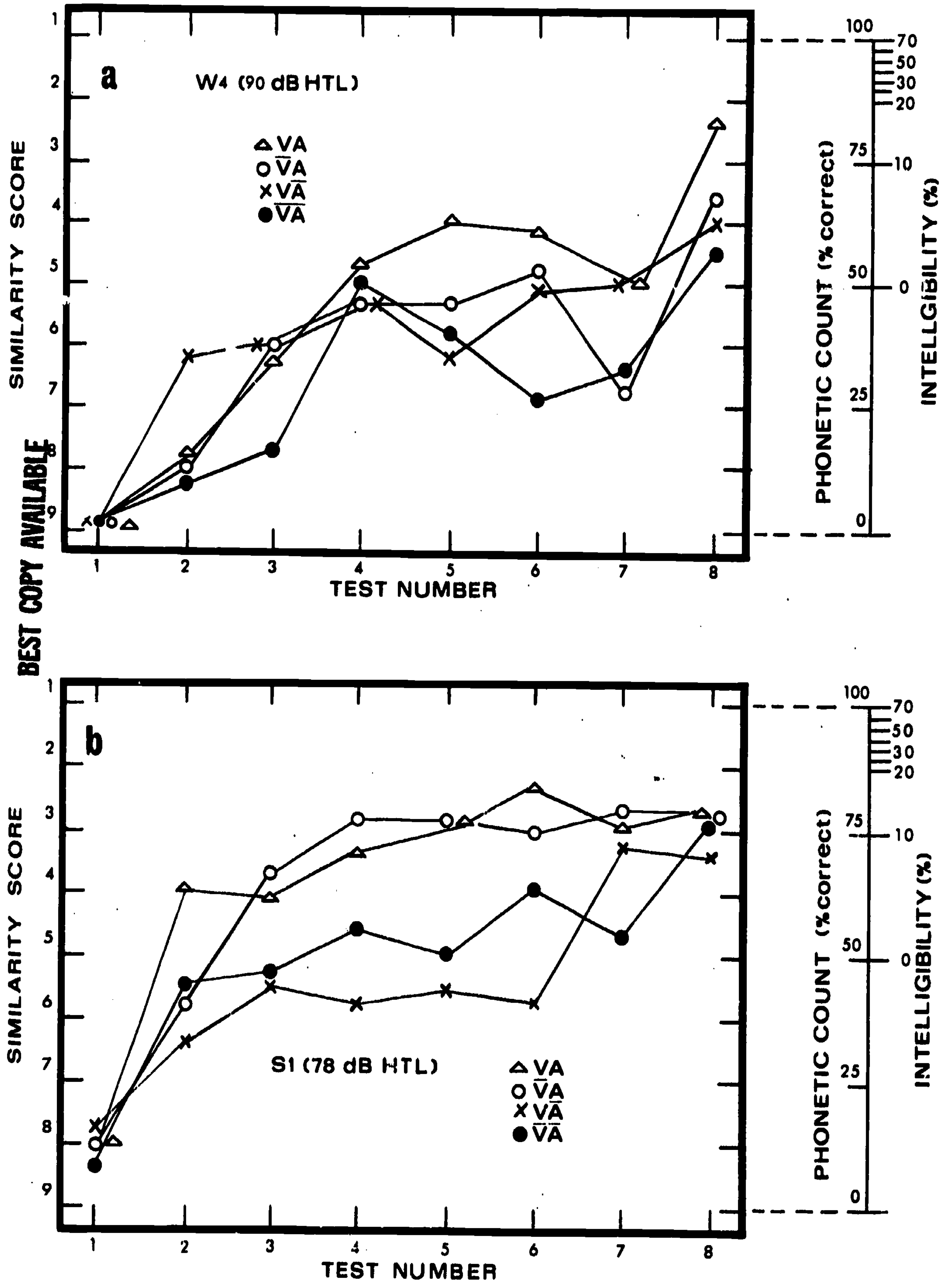


Figure 4: Mean Similarity Scores for Two Subjects (W4 & S1) as a Function of Four Test Conditions and the Relationship of these Scores to Phonetic Count and Intelligibility

To estimate the reliability of obtaining acoustic measures from the spectrograms, 25 words were randomly selected and analyzed the second time with the same procedure approximately six months after the first analysis. The Pearson product moment correlation coefficients (Ferguson, 1966) for the measures were as follows: 1) F_0 = 0.99; 2) contour = 0.87; 3) syllable match = 0.86; 4) voice duration = 0.95; and 5) latency = 0.92.

To estimate the range of the acoustic measures for normal-hearing children, four normal-hearing children with a similar age range (2 years, 8 months, to 6 years, 6 months) were evaluated by the same tester using the 15-word test (see Table 4). The test condition was without amplification and without visual clues ($\overline{VA}$). These speech samples were analyzed spectrographically using the same procedure described above. The mean values were: 1) F_0 = 289 Hz (range from 246 to 358 Hz); 2) contour = 89% correct (range from 80 to 100%); 3) syllable match = 95% correct (range 93 to 100%); 4) voice duration = 471 msec (range 382 to 602 msec); and 5) latency = 355 msec (range 220 to 525 msec). In addition, the oldest child (6 years, 6 months) achieved a mean similarity score of 1.2. With regard to the age of the children, F_0 was 358 Hz for the child of 2 years, 9 months, and it was 246 Hz for the child of 6 years, 6 months. For the other four measures, there was no observable relationship between age and the acoustic measures.

In addition, the same acoustic measures were obtained for the tester by analyzing 45 test words (three 15-word lists) that she used for evaluating the hearing-impaired children. The mean values for the tester were: 1) F_0 = 252 Hz, and 2) voice duration = 455 msec. The latter measure was used to compute the ratio of the voice duration between the tester and the child. The tester had a higher F_0 in a test situation than the 200 Hz that was measured for normal conversational speech (see Appendix E).

The acoustic measures for the Warren and Suvag groups are displayed in Tables 7-a and 7-b, respectively. This includes the five acoustic measures identified earlier, the similarity scores, and the results of the phonetic count which will be discussed later. Whenever possible these acoustic measures were obtained for every other test (e.g., Test 1, Test 3, and Test 5). It was not possible to analyze some samples because of an absent response or a poor signal-to-noise ratio on the tape. For each parameter (column), the mean value is listed for <u>all</u> words that were available for each test. The extreme right hand column identifies the number of words for each analysis of each subject. For example, for Tests 1, 3, and 5, W1 had a mean F_0 of 386, 369, and 449 Hz for 12, 15, and 15 words, respectively. The group means for each test and the overall mean is listed at the bottom of each table.

Because of an unequal number of words, the comparison between the Warren and Suvag groups as a function of tests was not feasible. As a result, a t-test (Ferguson, 1966) was used to compare 9 Warren subjects and 9 Suvag subjects at Test 3 for each of the above measures. The mean value for each subject was used for these analyses. For the measure of voice duration, the Suvag group (508 msec) had a significantly

Table 7-a. Acoustic measures, phonetic count, and similarity score for the Warren group for the testing condition of amplification without visual clues, ($\overline{VA}$).

BEST COPY AVAILABLE

Subject Number	Test	Mean Fundamental F_o (Hz)	Contour (% Correct)	Syllable Match (% Correct)	Voice Duration (msec)	Latency (msec)	Phonetic Count (% Correct)	Similarity Score	Number of Words
W1	1	386	80	63	277	901	8	8.3	12
	3	369	40	73	207	500	16	8.5	15
	5*	449	20	73	377	171	14	8.5	15
	Prob.	0.01**		0.61	0.01	0.58	0.32	0.31	
W2	1	314	0	80	765	1056	22	7.5	5
	3*	342	50	88	439	366	13	8.4	15
	5	298	10	77	653	723	23	7.8	15
	Prob.	0.17		0.62	0.21	0.29	0.81	0.27	
W3	1	359	10	89	349	513	41	5.6	14
	3*	38[illegible]	40	89	391	447	37	6.9	14
	Prob.	0.03**		1.0	0.54	0.54	0.52	0.14	
W4	1	248	60	85	278	1267	17	8.0	14
	3	229	80	82	321	508	46	5.2	14
	5*	224	70	89	324	559	61	4.6	14
	8	235	50	94	319	502	62	3.6	15
	Prob.	0.02**		0.22	0.57	0.00**	0.00**	0.00**	
W5	1	268	70	97	358	308	87	1.6	15
	3	256	100	100	359	243	97	1.4	15
	5*	255	90	100	460	377	100	1.1	15
	Prob.	0.01**		0.38	0.00**	0.01**	0.02**	0.00**	
W6	1	331	0	38	902	781	16	8.1	13
	3*	403	20	67	446	545	21	7.7	14
	Prob.	0.00**		0.01**	0.00**	0.20	0.74	0.67	
W7	1	664	0	51	730	621	20	8.2	14
	3*	372	50	77	362	521	37	7.4	10
	Prob.	0.00**		0.06	0.00**	0.84	0.12	0.55	
W8***									
W9***									
W10	3	310	0	72	445	508	20	7.6	12
W11	1	334	10	67	871	657	23	8.4	12
	3	306	20	64	562	507	40	7.5	6
	Prob.	0.57		0.61	0.12	0.52	0.71	0.33	
Test 1, Means, N=8		363	29	71	566	763	29	7.0	12
Test 3, Means, N=9		330	44	79	392	461	36	6.7	13
Test 5, Means, N=4		307	48	85	454	458	50	5.5	15
Overall Means		334	40	78	463	572	37	6.5	13

*Latest test used for calculation of analysis of variance
**$p < 0.05$
***No data available.

Table 7-b. Acoustic measures, phonetic count, and similarity score for the Suvag group for the testing condition of amplification without visual clues. ($\overline{V}A$).

BEST COPY AVAILABLE

Subject Number	Test	Mean Fundamental F_o (Hz)	Contour (% Correct)	Syllable Match (% Correct)	Voice Duration (msec)	Latency (msec)	Phonetic Count (% Correct)	Similarity Score	Number of Words
S1	1	568	40	74	351	909	12	7.9	15
	3	382	70	97	424	717	60	3.5	15
	5*	350	30	96	406	622	68	2.6	14
	8	330	80	100	339	621	86	2.2	14
	Prob.	0.00**		0.02**	0.14	0.02**	0.00**	0.00**	
S2	1	407	40	82	702	879	15	7.2	13
	3*	367	10	37	437	696	31	7.3	15
	5	320	40	93	[illegible]79	410	65	3.7	14
	Prob.	0.01**		0.80	0.01**	0.30	0.25	0.71	
S3	1	334	20	73	622	314	22	7.9	13
	3*	278	40	100	407	724	17	6.6	15
	5	242	80	82	326	5[illegible]3	23	7.1	14
	Prob.	0.02**		0.00**	0.03**	0.01**	0.85	0.06	
S4	1	374	90	97	469	597	66	4.4	15
	3*	337	50	100	555	[illegible]33	69	3.6	15
	5	334	60	96	424	553	85	2.6	14
	Prob.	0.00**		0.66	0.04**	0.53	0.73	0.06	
S5	1	229	0	75	829	247	8	8.0	14
	3	295	10	83	500	505	18	7.7	14
	5*	265	30	77	381	702	24	7.0	11
	8	281	10	89	579	565	36	6.6	15
	Prob.	0.00**		0.60	0.01**	0.04**	0.13	0.10	
S6	1	650	10	73	1452	572	8	9.0	10
	3*	348	20	46	453	621	23	8.0	4
	5	438	40	83	605	978	21	7.3	15
	Prob.	0.24		0.40	0.11	0.45	0.50	0.50	
S7	1	309	0	71	538	1476	13	7.8	4
	3*	284	30	63	599	623	19	8.4	7
	Prob.	0.69		1.00	0.98	0.53	0.50	0.56	
S8	3	334	10	74	750	878	27	3.8	15
S9	1	310	70	[illegible]3	456	379	69	4.3	14
	3*	300	90	97	447	403	74	3.2	15
	Prob.	0.04**		0.66	0.88	0.66	0.69	0.10	
Test 1, Means, N=8		398	30	80	677	672	27	7.1	12
Test 3, Means, N=9		325	40	83	508	651	38	5.8	13
Test 5, Means, N=6		325	50	88	437	631	48	5.1	14
Overall Means		347	40	84	541	648	38	5.9	13

*Latest test used for calculation of analysis of variance
**$p < 0.05$

longer duration ($t = 2.34$; $df = 16$; $p < 0.05$) than did the Warren group (392 msec). As a comparison, the normal-hearing children (471 msec) had a greater voice duration than the tester (455 msec). The Suvag group appeared to be in closer approximation to the normal-hearing children than did the Warren group. This analysis was performed prior to computing the ratio for this measure.

For the measure of latency, the Warren group (461 msec) had a significantly shorter latency ($t = 3.41$; $df = 16$; $p < 0.01$) than the Suvag group (651 msec). The latency for the Warren group is closer to the latency of the normal-hearing children (355 msec). The remaining between-group comparisons for the acoustic measures were not significant.

To have some estimate of the significant changes as a function of tests, a one-factor analysis of variance (Winer, 1971) was used to analyze the common words for each subject. The mean for each subject was based on the same words across Tests 1 and 3, and 5, if available. The probability associated with each analysis for each subject is displayed in Tables 7-a and 7-b. For some subjects additional measures for Tests 5 and 8 were obtained too late in the project for inclusion in the calculation of the analysis of variance. The significant differences ($p \leq 0.05$) are identified by two asterisks (**).

For example, W4 displayed a significant consistent decrease ($F = 4.4$; $df = 2,22$; $p < 0.03$) in F_0 from Test 1 (248 Hz), to Test 3 (229 Hz), to Test 5 (224 Hz). Some changes were difficult to interpret because some subjects decreased and then increased; however, some general trends were observed.

For subjects with an F_0 within the typical range for normal-hearing children, there was not a great change as a function of the tests. However, the children who had an abnormally high F_0 at Test 1 generally lowered their F_0 to approximately the normal range for Test 3.

For intonational contour, there was an increase in the percent correct from Test 1 to Test 3. Some of the better subjects reached the normal range at Test 5 and Test 8.

The syllable match for most of the subjects was far below the normal range at Test 1, but it improved with training. Some of the subjects reached the normal range at Test 5.

The voice duration decreased as the number of tests increased. Time did not allow for the conversion of these data to ratios.

For most of the subjects, the latency measure at Test 1 was much longer than the latency for normal-hearing children, but it decreased with training. Even with the decrease, most of the subjects had longer latencies than normal-hearing children.

Because the previous acoustical analyses at the subject level did not produce any significant trends, a more detailed analysis was performed at the word level. The words were grouped in equal intervals of 0.5 based on the similarity score that had been previously assigned. The mean and standard deviation were computed for each individual based on the number of words that were available. These values are displayed in Table 8. For example, the interval of 1.0 to 1.4 had a mean fundamental frequency of 285.4 Hz for 62 words that were spoken by seven subjects. The voice duration is expressed as a ratio between the duration of the child's response and the teacher's stimulus. A ratio larger than 1 represented responses that were longer than the teacher's stimuli, and the ratios smaller than 1 are for responses that were shorter than the stimuli. The measures on phonetic count will be discussed later in this report.

The last column on the right identifies the number of words and subjects contributing to each interval on the scale. The number of words varied between 15 and 131; the number of subjects varied between 6 and 16.

The mean and standard deviations from Table 8 are plotted in Figures 5-a to 5-e as a function of the similarity score. The means are identified by crosses, and the standard deviations by open circles. As the speech production improved from 9 to 1 in similarity score (see Figure 5-a), F_0 decreased 76 Hz, from 362 Hz to 285 Hz. It appears that both training and age contributed to this improvement; however, if young deaf children do not receive the proper training, it is observed that the F_0 remains higher than that of normal-hearing children of similar ages.

Intonational contour displayed a 51% improvement (25 to 76%) from 9 to 1 on the scale (see Figure 5-b). The 76% level was slightly below the range (80 to 100%) for normal-hearing children.

Syllable match improved 24% (74 to 98%) and reached the range for normal-hearing children (93 to 100%) (see Figure 5-c). It appeared that improvement in syllable match was much easier for the children to achieve than improvement in intonational contour.

The mean ratio for the voice duration (Figure 5-d) showed a gradual decrease as similarity score decreased. This means the child's voice duration changed from a duration that was longer than the teacher's stimuli to a duration that was either equal to or less than the duration of the teacher's stimuli.

The mean values for latency (Figure 5-e) tended to decrease, but displayed great variability for adjacent intervals on the scale. The shortest latency (449 msec) for a similarity score of 1 was less than the longest duration (525 msec) for the normal-hearing children, but longer than the mean normal latency (355 msec).

There was a noticeable decrease in standard deviation for F_0, syllable match, voice duration, and latency as similarity score decreased. This indicated less variability among the words for the lower portion of the similarity scale. There was no standard deviation for intonational contour because it was evaluated on a 2-point scale (see Appendix J).

Table 8. Means and standard deviations of the acoustic measures and the phonetic count for each 0.5 interval on the similarity scale

BEST COPY AVAILABLE

Similarity Score	Fundamental Frequency F_o (Hz)		Syllable Match (% Correct)		Contour (% Correct	Voice Duration Child/Tester		Latency (msec)		Phonetic Count (% Correct)		Number of Words	Number of Subjects
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD		
1.0–1.4	285	51	98	11	76	0.85	0.39	449	282	93.6	16.7	62	7
1.5–1.9	316	61	99	8	72	0.77	0.20	492	250	88.1	17.5	[illegible]	9
2.0–2.4	308	54	98	11	68	0.99	0.40	641	488	71.6	30.6	22	7
2.5–2.9	300	60	89	20	46	0.90	0.39	483	261	76.2	26.9	22	9
3.0–3.4	326	39	94	15	67	1.02	0.25	546	289	80.6	20.5	15	6
3.5–3.9	311	56	93	17	44	0.87	0.24	591	157	52.8	25.8	18	9
4.0–4.4	345	52	94	16	33	0.98	0.38	705	391	59.3	24.1	18	11
4.5–4.9	301	70	88	21	48	0.86	0.36	509	160	47.9	24.5	23	9
5.0–5.4	318	87	95	15	45	0.80	0.34	533	191	30.5	33.7	20	9
5.5–5.9	284	70	88	22	44	0.86	0.17	576	231	45.3	30.5	16	9
6.0–6.4	353	94	80	23	42	1.01	0.46	646	288	41.7	30.8	26	14
6.5–6.9	329	64	80	25	42	0.79	0.34	613	270	33.7	23.9	24	11
7.0–7.4	362	96	73	27	43	0.87	0.49	496	343	25.4	21.4	42	15
7.5–7.9	369	131	86	25	32	1.04	0.50	619	302	25.3	29.6	31	12
8.0–8.4	343	94	78	26	21	1.07	0.70	640	452	15.2	19.7	57	14
8.5–9.0	362	113	74	26	25	1.14	0.85	654	797	9.4	16.9	131	16

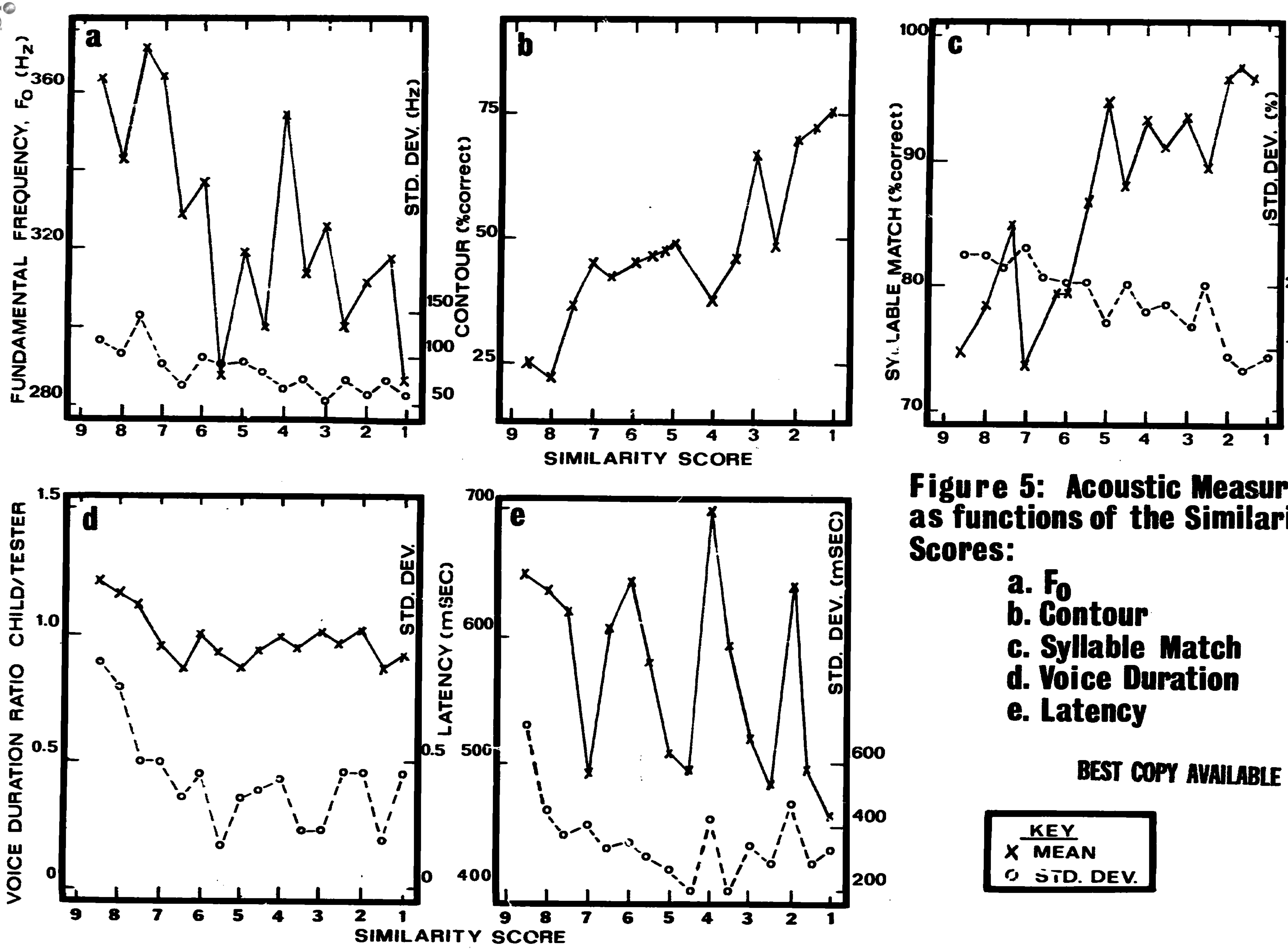


Figure 5: Acoustic Measures as functions of the Similarity Scores:

a. F_0
b. Contour
c. Syllable Match
d. Voice Duration
e. Latency

BEST COPY AVAILABLE

KEY
X MEAN
O STD. DEV.

A Pearson product moment correlation coefficient (Ferguson, 1966, p. 111) was computed for the above acoustic measures relative to similarity score ($N = 16$). The coefficients were:

1) $F_0 = +0.72$
2) Intonational contour = -0.84
3) Syllable match = -0.87
4) Voice Duration = +0.45
5) Latency = +0.48

The first three acoustic measures had a higher correlation with the similarity scores than was detected for voice duration and latency.

These results indicate that acoustic measures can be a valuable tool in evaluating changes in speech production of deaf children across the continuum of the sub-intelligible to the intelligible range. Because several different parameters can be studied simultaneously, these measures should provide more insight into how these suprasegmental features interact to improve the speech of deaf children.

C. Phonetic Count

In order to gain additional insight a phonetic transcription of each speech sample (stimulus and response) was written on each corresponding spectrogram that was used for the acoustical analyses. This transcription was made by a person highly trained in listening to the speech of the deaf. Each transcription was compared to the stimulus to determine the number of phonemes that were correctly imitated by the child. For each subject, a phonetic count was computed for the condition of amplification without visual clues ($\overline{V}A$) (see Tables 7-a and 7-b).

To estimate the reliability of the transcription, 25 samples that had been analyzed earlier were randomly selected and re-analyzed with the same procedure and the same listener. The Pearson product moment correlation coefficient (Ferguson, 1966) was 0.95 which suggested a high degree of reliability for the phonetic count.

For each subject, a one-factor analysis of variance (Winer, 1971) was computed to determine the significance of differences which occurred ($p \leq 0.05$). Only subjects W4, W5, and S1 demonstrated a significant improvement (increase in phonetic count) for Tests 1, 3, and 5. The same statistical procedure was used for the similarity scores (see Tables 7-a and 7-b) that corresponded to the samples from the spectrographic analyses. These same three subjects (W4, W5, and S1) demonstrated significant changes in similarity scores.

Next, a phonetic count was computed at the word level according to equal intervals (0.5) on the similarity scale. The mean and standard deviation for each interval are displayed in Table 8 and also plotted as open circles in Figure 6. These data on phonetic count can be approximated by the straight line with the 45° slope. A Pearson correlation coefficient (Ferguson, 1966, p. 111) was computed between the phonetic count and similarity score ($N = 16$). The coefficient was -0.96.

This indicated that a similarity score and a phonetic count were linearly related to each other. The improvements in both measures were indications that the training method was successful and that both are useful in evaluating the unintelligible speech patterns of deaf children. The phonetic count is less time-consuming than obtaining similarity scores; therefore, it might be more feasible for evaluating progress.

D. Intelligibility

As mentioned previously, it was not possible to utilize the standard intelligibility tests because many of the utterances from the deaf children were unintelligible. However, with significant changes in similarity score for both groups, it became necessary to identify the portion of the similarity scale that contributed to intelligible speech.

The speech samples recorded for the $\overline{V}A$ condition were separated from the stimuli, randomized, dubbed on a separate tape, and judged by a panel of 22 normal-hearing college students. Words produced by 19 subjects were used for this analysis. The listeners wrote a word for each response of a child. Table 9 displays the intelligibility score in percent correct for each interval. The intervals were as follows: 1) 0.2 for similarity scores between 1.0 and 3.0; 2) 0.5 for scores between 3.0 and 5.0; and 3) 1.0 for scores between 5.0 and 9.0. For each interval on the scale, all 15 test words were represented with a random selection across subjects.

Table 9. Mean of intelligibility scores for intervals on the similarity scale.

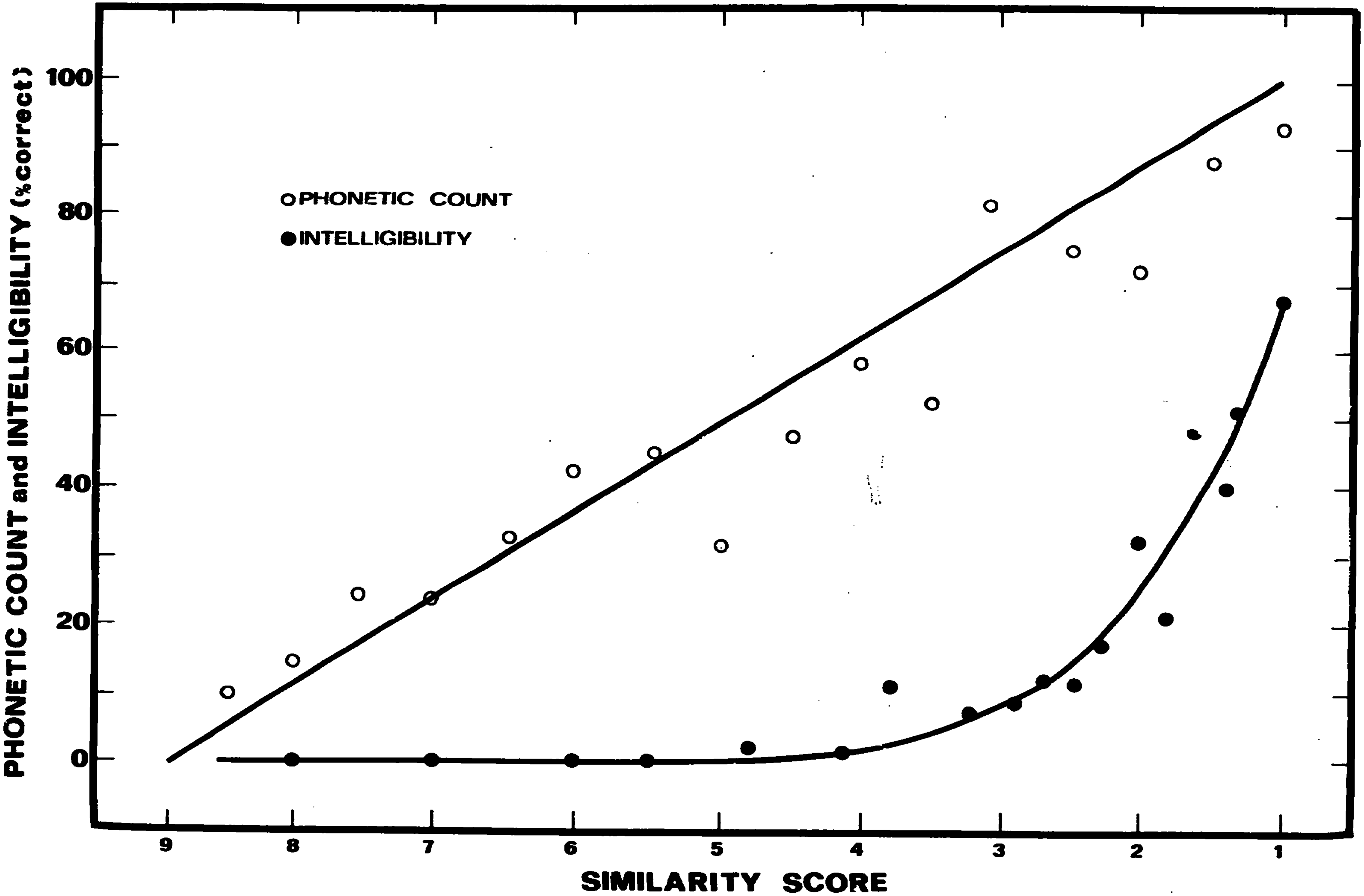
Similarity Score	1.0-1.1	1.2-1.3	1.4-1.5	1.6-1.7	1.8-1.9	2.0-2.1	2.2-2.3	2.4-2.5	2.6-2.7	2.8-2.9
Intelligibility (% Correct)	67	51	40	48	24	33	16	15	14	9

Similarity Score		3.0-3.4	3.5-3.9	4.0-4.4	4.5-4.9	5.0-5.9	6.0-6.9	7.0-7.9	8.0-9.0	
Intelligibility (% Correct)		6	11	3	2	0.9	0	0	0	

The mean intelligibility scores from Table 9 are plotted graphically in Figure 6 as the solid circles. The exponential line appears to be the best approximation of these data points. Between 9 and 5 on the similarity scale, the scores were 0% correct in intelligibility. From 5 to 1 on the scale, the scores increased up to approximately 70% correct for a scale value of 1. As a result, the region between 9 and 5 on the similarity scale was identified as the sub-intelligible range and the area between 5 and 1 as the intelligible range. The non-linear growth of the intelligibility scores limits the usefulness of the measure for evaluating the progress.

Figure 6: Phonetic Count and Intelligibility as Functions of the Similarity Scores

BEST COPY AVAILABLE



The data from phonetic count and intelligibility scores that are displayed in Figure 6 were used to interpret the changes that occurred for subjects W4 and S1 discussed earlier in this report (see Figure 4). At Test 8, both subjects have greater than 50% of the phonemes correct and achieve an intelligibility score between 1 and 19% correct, depending on the test conditions. This degree of improvement is rather remarkable when one considers that single word test items are a stringent measure of speech perception and production for young deaf children.

E. Additional Measures and Analyses

In search of the best measure of progress for these deaf children, several other approaches were undertaken. They will be mentioned briefly at this point.

1) Analysis of the type of responses. To evaluate the type of response, the following numbers were assigned to all responses according to the following three criteria: "1" - a response by the child that was recorded and judged by the listeners; "2" - a response that was so poor that it was excluded from the randomized sample and was previously assigned a rating of 9; and "3" - an absent response that was in the previous analysis automatically assigned a rating of 9. These responses were ordered from the most desirable ("1") to the least desirable ("3"). For each subject, the criterion measure was the mean "response level" of the 15-word test for each testing condition.

A four-factor analysis of variance (Winer, 1971) with repeated measures on three factors was used to analyze this criterion measure for twenty subjects for Tests 1, 2, and 3. Although there was a significant decrease (improvement) ($F = 7.57$; $df = 2,36$; $p < 0.002$) in response level for three tests (Test 1 = 1.83; Test 2 = 1.47; and Test 3 = 1.35), the response level of the Warren group (1.63) was not significantly different ($F = 1.37$; $df = 1,18$; $p < 0.56$) from that of the Suvag group (1.45). The other main factors and interactions were not significant. As with the analysis for the similarity scores, the results were significant for tests, but not significant for the between group measures.

2) Hearing aid evaluation. As indicated earlier, each subject was fitted with a Mini-Suvag or a Zenith Vocalizer II hearing aid according to the group to which he had been assigned. The Interim Report (Asp, 1972, pp. 22-25) describes the hours of use for each type of aid. Since speech-discrimination testing could not be accomplished with these young deaf children, the aids were evaluated subjectively. Upon completion of the project, four children who had previously worn Zenith hearing aids were fitted with Mini-Suvag aids. The children's parents and teachers were asked to fill out a questionnaire. Questions were asked regarding differences in behavior, the number, type, and quality of vocalizations, difficulty in adjusting to the hearing aid, etc. Although the evaluations of the hearing aids were all slightly different, one response which consistently appeared on the questionnaire was that all parents thought, for whatever reasons, that the Mini-Suvag aid was helping their child, and none preferred that the child be refitted with the previously-worn Zenith hearing aid.

3) <u>Half-harmonics</u>. In analyzing spectrograms for the acoustic measures, half-harmonics (see Appendix J) were observed for some subjects and some of the speech samples. Table 10 lists only the subjects who had half-harmonics and provides the percentage of half-harmonics for each 15-word test. The majority of subjects decreased in observable half-harmonics as test number increased.

Table 10. Percentage of responses for which half-harmonics were present across tests.

Subject	Test 1	Test 2	Test 3	Test 4
W1	67	0	0	-
W2	60	33	-	-
W4	14	7	0	13
W6	0	7	-	-
W7	0	50	-	-
S1	67	7	0	29
S4	0	13	-	-
S6	20	0	-	-

- = Data not available

FILTERED SPEECH TESTING

Verbo-tonal audiometry includes a battery of tests that are used for diagnosis of hearing impairment. One of these tests uses nonsense syllables as stimuli; these nonsense syllables are identified as logatomes. The logatomes are: /bru, mu, bu, vo, la, ke, ʃi, si/. The rationale and a detailed description of the stimuli will not be included, but it can be found in the Interim Report (Asp, 1972). The criterion for selecting the logatomes is the pitch of the consonant(s) and vowel. The logatomes range from low to high pitch. For example, /bru/ has the lowest pitch and /si/ the highest pitch.

For the present study, the logatomes were recorded on magnetic tape by a male voice. Each logatome was spoken twice in succession (e.g., /si-si/), so the stimuli would have the natural rhythm of speech. These pairs of logatomes were re-recorded in two forms: unfiltered and filtered. The filtered logatomes were passed through octave bandpass filters according to the optimal octaves recommended by Guberina (1964). The pair /si-si/ was passed through two different filters producing two independent stimuli. These stimuli ranged from the lowest octave band of 50 to 100 Hz to the highest octave band of 6.4 to 12.8 kHz. Only the filtered logatomes were used in this study.

Standard techniques for play audiometry were used to obtain detection thresholds for both pure tones and the filtered logatomes. First, pure-tone testing was accomplished at least three times a year with a standard portable audiometer (Beltone, model 10-D). The children were well-conditioned and the test usually could be completed within a single session. Tables 2-a and 2-b display the pure-tone thresholds for the Warren and Suvag subjects, respectively.

After pure-tone thresholds were established, filtered speech testing was undertaken in 1-dB steps using the filtered logatomes. Since the children were less familiar with this test, testing usually was completed only after two or three sessions. Tables 11-a and 11-b display the detection thresholds for the filtered logatomes for the Warren and Suvag subjects, respectively. The results of Tables 11-a and 11-b were compared with the pure-tone thresholds presented in Tables 2-a and 2-b for the two groups of subjects. The test signals for the pure-tone audiometry were seven standard test signals ranging from 125 to 8000 Hz. The corresponding seven filtered logatomes are identified as 2A through 8A (see Table 11).

The detection thresholds in Tables 2 and 11 are for both right and left ears. An asterisk identifies the better ear for each subject relative to the three-frequency average of 0.5, 1, and 2 kHz. The mean values for each group for both the better and poorer ears are displayed at the bottom of Table 2; Table 11 includes the group means for the right and left ear.

To evaluate possible differences between the Warren and Suvag groups, both the better ear thresholds and the composite of better thresholds for pure-tone and filtered logatomes were used. For the composite of better thresholds, the best threshold was selected for each test signal regardless of the ear. Table 12 displays the group means and the differences between the group means for the better ear and better threshold. An analysis of variance (Winer, 1971) was used to compare the two groups. The Warren group obtained a poorer mean threshold in each case. The differences between the groups were greater for the pure-tone signals (better ear = 2.9 dB, better threshold = 3.1 dB) than for filtered logatomes (better ear = 0.8 dB, better threshold = 1.6 dB); however, the F-ratios were not statistically significant. As a result, in the following analyses, the 20 subjects are treated as one group.

Table 12. Mean detection thresholds in decibels (ISO 1964) for Warren and Suvag groups as a function of pure-tone and filtered logatomes for better-ear thresholds and composite of better thresholds.

	Pure-tone (Air Conduction)			Filtered logatomes (Air conduction)		
	Warren N = 11	Suvag N = 9	Difference	Warren N = 11	Suvag N = 9	Difference
Better Ear	85.3	82.4	2.9	87.0	86.2	0.8
Better Threshold	84.9	81.7	3.2	86.3	84.7	1.6

There was no significant difference between the two tests (pure-tone and filtered logatome) for either better ear measures ($F = 3.57$; $df = 1,18$; $p < 0.07$) or the composite of better thresholds ($F = 2.69$; $df = 1,18$; $p < 0.12$). This indicated that the pure tones and filtered logatomes resulted in similar mean thresholds for both groups.

Table 11-a. Verbo-tonal filtered air-conduction detection thresholds in dB (ISO 1964) for the Warren group

Subject No.	Ear	Octave Test Bands in Hz for Each Logatome								
		bru-bru 50-100 1B	mu-mu 75-150 2A	bu-bu 150-300 3A	vo-vo 300-600 4A	la-la 600-1200 5A	ke-ke 1200-2400 6A	ʃi-ʃi 2400-4800 7A	si-si 4800-9600 8A	si-si 6400-12800 8B
W1	R*	52	59	68	66	86	98	102	107	99
	L	65	73	81	90	91	87	101	119+	107+
W2	R*	66	70	82	96	99	75	97	109	100
	L	77	83	95	96	107	95	109	110	107+
W3	R*	41	62	72	85	86	67	61	74	71
	L	56	65	85	86	81	73	62	81	75
W4	R*	68	75	82	78	76	95	89	111	99
	L	70	79	88	98	111	84	90	103	100
W5	R	50	46	57	62	71	60	54	67	67
	L*	47	50	50	55	57	62	56	59	57
W6	R*	82+	89	92	110	127	118	114	119+	107+
	L	82+	95	105	124	127	110	112	119+	107+
W7	R	70	83	88	108	111	116	115	119+	107+
	L*	72	99	114	116	128	122	124+	119+	107+
W8	R	63	81	103	95	94	82	72	89	94
	L*	65	80	91	89	94	83	69	79	71
W9	R	82+	85	102	97	107	107	114	119+	107+
	L*	62	60	80	90	89	85	94	99	100
W10	R*	76	83	95	95	108	102	104	116	107+
	L	74	80	95	96	107	104	109	119+	107+
W11	R*	47	68	81	91	95	107	114	119+	107+
	L	63	72	82	91	100	105	115	119+	107+
MEAN	R =	61	70	80	88	93	88	92	>100	>93
MEAN	L =	69	79	92	96	103	97	>97	>107	>99

* = Better at 4A, 5A, and 6A

> = Greater than

Table 11-b. Verbo-tonal filtered air-conduction detection thresholds in dB (ISO 1964) for the Suvag group

Subject No.	Ear	Octave Test Bands in Hz for Each Logatome bru-bru 50-100 1B	mu-mu 75-150 2A	bu-bu 150-300 3A	vo-vo 300-600 4A	la-la 600-1200 5A	ke-ke 1200-2400 6A	ʃi-ʃi 2400-4800 7A	si-si 4800-9600 8A	si-si 6400-12800 8B
S1	R	77	70	82	107	119	100	99	117	107+
	L*	52	64	67	74	85	64	56	70	67
S2	R	79	90	102	110	112	83	81	109	97
	L*	64	69	98	88	97	92	96	119+	107+
S3	R	65	77	86	97	106	108	116	119+	107+
	L*	60	70	64	82	99	93	104	106	99
S4	R	60	75	90	100	100	104	104	114	107+
	L*	62	70	77	90	77	60	46	60	54
S5	R*	61	63	77	96	109	95	100	102	97
	L	70	81	93	109	119	105	114	119+	107+
S6	R	77	83	97	102	111	110	114	119+	107+
	L*	79	84	97	101	112	105	106	119+	107+
S7 **	R	77	80	92	120	131	115	114	119+	107+
	L*	77	85	92	102	112	110	99	119	107+
S8	R	67	80	89	108	111	95	108	119+	107+
	L*	69	80	91	104	106	89	101	119+	107+
S9	R*	25	42	48	50	55	73	92	96	79
	L	55	69	67	63	64	85	90	119+	107+
Mean	R =	60	69	79	87	95	86	87	>100	>90
Mean	L =	70	79	89	102	108	102	106	>110	>107

* = Better ear at 4A, 5A, and 6A

** = Difficult to test

> = Greater than

For the seven test frequencies, there was a significant difference for both the better ear ($F = 19.31$; df = 6,108; $p < 0.0001$) and the composite of better thresholds ($F = 20.66$; df = 6,108; $p < 0.0001$). The interaction of tests and frequencies was significant for both the better ear ($F = 9.81$; df = 6,108; $p < 0.0001$) and better thresholds ($F = 11.22$; df = 6,108; $p < 0.0001$).

The better thresholds are listed in Table 13 for both tests as a function of the seven test signals. The right-hand column identifies the difference in dB for the two test signals. Pure-tone thresholds are consistently better than the filtered logatome thresholds at 125 Hz (5.2 dB), 250 Hz (8.5 dB), and 8000 Hz (9.7 dB). On the other hand, filtered logatome thresholds are better than the pure-tone thresholds at 2000 Hz (6.2 dB) and 4000 Hz (6.3 dB). Although the differences of 6 to 10 dB would not normally be significant for individual comparisons, the differences across groups appear to be worthy of consideration. For the low-frequency test signals, ambient noise may create a more difficult listening situation for filtered logatomes than for pure tones. For the higher frequencies, the 1200 to 2400 Hz and 2400 to 4800 Hz octave bands provide a wider band of energy than the corresponding pure-tone signals. For the 6400 to 12,800 Hz band, the signal-to-noise ratio was the poorest because of the small amount of speech energy present in the logatome /si-si/ in this frequency range. This may have contributed to the poorer threshold (9.7 dB) for this logatome.

Table 13. Mean detection thresholds in decibels (ISO 1964) for 20 subjects for pure tones and filtered logatomes based on the composite of better thresholds.

Pure tones		Filtered logatomes		Difference between thresholds (dB)
Frequency (Hz)	Threshold (dB)	Frequency range (Hz)	Threshold (dB)	
125	64.5	75-150	69.7	-5.2
250	71.0	150-300	79.5	-8.5
500	86.3	300-600	87.5	-1.2
1 K	90.8	600-1.2 K	93.7	-2.9
2 K	93.5	1.2-2.4 K	87.3	+6.2
4 K	96.0	2.4-4.8 K	89.7	+6.3
8 K	82.3	6.4-12.8 K	92.0	-9.7

To estimate the correlation between thresholds for pure tones and filtered logatomes, a Pearson product moment correlation (Ferguson, 1966) was used. The correlation coefficients from the lowest to the highest frequency test signals were 0.61, 0.77, 0.77, 0.78, 0.68, 9.91, and 0.81, respectively. In general, the correlation coefficient increased as the test frequency increased. The filtered logatomes and the pure tones measure similar detection thresholds, so the filtered logatomes do not provide any additional information for groups of young deaf children.

DISCUSSION

The results of this study did not show a difference between the speech reception/production skills of two groups of young deaf children when frequencies below 200 Hz were included as part of the daily training program for one of the groups. Related studies by other investigators have shown results that conflict with each other and with this study. For example, some studies cite the benefits of low-frequency amplification (Ling, 1963, 1964a, 1964b, 1966; Briskey and Sinclair, 1966; Briskey, Garrison, Owsley, and Sinclair, 1967; Leckie and Ling, 1968), while others infer a deleterious (masking) effect of low-frequency amplification (e.g., Martin and Pickett, 1970) upon the speech reception of the hearing impaired.

The two major factors to consider in comparing the results of these studies are the age of the subjects and the effects of auditory training. In the case of the latter, most of the studies that cited the benefits or the deleterious effects of low-frequency amplification did not specifically train the subjects through the amplification systems beyond the regular testing period. The main goal was to detect changes in speech reception without including formal training. In the present study, the emphasis was on the effect of auditory training on the speech reception/production skills of two groups trained with two different amplifying systems. Thus, the amount and the type of auditory training may have minimized the differences between these two different types of amplification.

With respect to the age of the subjects, Ling (1963), for example, used children of school age, Martin and Pickett (1970) used college students, and none of the other studies used preschool children as the present study did. The low frequencies may be very important for developing the suprasegmental aspects of speech during the early formation of the child's perceptual and production skills, as hypothesized by Guberina (1964). Then, after the child has developed the basic speech skills, low frequencies may cause a masking effect (such as found by Martin and Pickett, 1970) on the perception of the segmental aspects of the speech signal.

Although our study demonstrated neither an advantage nor a disadvantage from the use of low-frequency amplification, this does not mean that a difference does not exist and that a low-frequency response may not be beneficial for some deaf children. Low-frequency amplification did <u>not</u> produce deleterious effects upon the children's performance when group means were evaluated, and no negative effects were reported for individual subjects.

It probably is naive to think that one frequency response can be appropriate for all young deaf children. The degree of loss, the age, and the perceptual skills the child has achieved must all be considered in the selection of the appropriate frequency response. Further, the type of frequency response may change as the child improves in perceptual skills. Thus, the most appropriate frequency response for hearing-impaired children probably should be selected and re-evaluated on an individual basis rather than on a group basis.

For both groups of deaf children, the results of this study indicate a significant improvement in the speech reception/production abilities during the preschool application of Verbo-tonal instruction. These results demonstrate that children with severe hearing impairments can be trained auditorily by an effective aural/oral method of instruction, such as the Verbo-tonal method. The consistent use of the appropriate type and level of amplification in conjunction with an effective instructional method is the ideal situation. These factors are not always included in training programs. For example, many programs use amplification as part of the training procedure and, on the basis of this use, identify the program as having an acoustical emphasis; however, most of these programs do not train the child to develop sophisticated auditory perception so that he might increase his chances of having normal perception and production of speech. Each deaf child should receive this training, so that he will have the opportunity to develop these skills.

The results also revealed that filtered-speech audiometric testing of young deaf children provides similar information to that of pure-tone audiometry, although the interaction of tests and frequency was significant. In general, the use of filtered logatomes required a longer testing time than the use of pure tones. One audiologist felt that children responded to filtered logatomes more easily, while the other two audiologists felt that the reverse was true. Based on our observations, it appears that filtered logatomes do not provide any additional information for differentiating between groups of young deaf children. However, in another study that we have undertaken with older children and adults, we find both filtered and non-filtered logatomes to be extremely beneficial in determining the appropriate frequency response for auditory training and hearing-aid evaluation. Thus, filtered-speech testing appears to be more useful with subjects who have developed a higher level of perceptual skills.

The new testing measures developed during this project are useful for evaluating progress in speech production from the sub-intelligible to the intelligible level. These measures include: a) the similarity scale, b) the phonetic count, and c) a battery of acoustic measures which includes fundamental frequency, intonational contour, syllable match, voice duration, and latency. Each is capable of measuring the entire continuum from the sub-intelligible to the intelligible level. Both the similarity score and the phonetic count provide a single measure that is easier to interpret than the acoustic measures. Of these two, the phonetic count is the more practical measure, for it is easier and less time-consuming to obtain. On the other hand, the acoustic measures provide more insight into the suprasegmental development of the speech because several parameters can be studied simultaneously.

A comparison of the five acoustic measures with the similarity score revealed that fundamental frequency, intonational contour, and syllable match come closer to the model provided by the normal-hearing speaker as the child improves in similarity score. The changes in voice duration and latency were more difficult to interpret. There is a need for more research on the acoustic measures to determine which measures most consistently detect changes in the suprasegmental aspects of speech development. This information is necessary in order to provide guidelines for teachers who are attempting to develop the suprasegmental aspects of the child's speech. The level to which the suprasegmentals develop probably determines the intelligibility level which a child will achieve.

This study demonstrated that conventional measures of intelligibility are not appropriate for evaluating the speech production of young deaf children because most of these children need a considerable amount of training before their speech is intelligible. The intelligibility scale has a non-linear growth and is not sensitive in the sub-intelligible range; therefore, it is not appropriate for evaluating most young deaf children. One can easily observe this phenomenon. For example, when a classroom teacher proudly refers to a deaf child who has made progress in speech production, the observer is usually puzzled because it is difficult to understand the child. One can only assume that the progress which the teacher is referring to is restricted to the sub-intelligible range, and therefore not detectable by a judgment of the intelligibility of the child's speech. This is an example of the importance of developing measures that are sensitive to changes in the sub-intelligible range of the child's speech. These measures are basic to establishing valid criteria for evaluating speech improvement in young deaf children.

In viewing the results of this study, one should be aware of the extreme difficulty of evaluating the speech communication abilities of very young deaf children. Among the more prominent problems are matching groups, controlling all variables, and implementing a formal testing procedure. For example, a preferred procedure for evaluating a child is to use a pre-recorded familiar voice (e.g., a teacher) and have the child make a written response (Erber, 1971b). This is not feasible for very young deaf children.

This study used single-word responses from the children because they were easier to evoke and analyze. These samples possibly were not of sufficient length to demonstrate differences in rhythm and intonation that might distinguish between types of amplification. The relationship between low frequencies and the development of rhythm and intonation is a major hypothesis of Guberina (1964), and it may not have been tested adequately in this study. For a future study, a test should be developed that uses a multi-word response, so that the suprasegmental aspects of speech can be adequately evaluated.

A detailed description of the Verbo-tonal method was not included in this report because the purpose of this study was the comparison of amplification systems and not an evaluation of instructional methods. However, Appendix F describes the classroom procedures used in the study and Appendix H lists video-tapes, articles that have been translated, and convention papers that pertain to this method. In addition, Appendix K is the curriculum guide that was written for the project to describe the use of the Verbo-tonal method with young deaf children. There is a need for further studies to gather information regarding the effectiveness of all the aspects of the Verbo-tonal method, e.g., vibro-tactile stimulation, rhythmical activities, etc., for this method produced a significant improvement in the speech productions of these deaf children. For example, vibro-tactile stimulation seems to be an effective tool for lowering the pitch of the child's voice. Some aspects of the Verbo-tonal method may be similarly effective and others may not. Each should be evaluated.

During this project, a remote FM system was developed for monitoring the speech of deaf children outside the classroom. The Interim Report (Asp, 1972) described the system and reported some of the measurements that had been obtained for vocalization rate and duration. Although these data were obtained, there was not a sufficient amount to complete a statistical analysis for group comparisons. However, the potential value of this technique for monitoring the speech behavior of deaf children should be explored.

This study was complex and time-consuming because the children were, among other things, young, handicapped, difficult to train, and especially difficult to test with some degree of reliability and validity. Because of the new measures developed for this study, it was possible to obtain meaningful data to describe speech improvement, especially in its suprasegmental aspects. More studies of this type are needed in order to provide classroom teachers with the necessary guidelines for habilitating deaf children so that they develop the necessary communication skills to function independently in the mainstream of our society.

CONCLUSIONS

1. There was no statistical difference between the speech reception/production abilities of two groups of young deaf children when frequencies below approximately 200 Hz are included as part of their daily training program. The results demonstrated neither an advantage nor a disadvantage from the use of low-frequency amplification. Although frequency response did not have an effect for groups of children, this does not mean that a low-frequency response is not important for some deaf children. Each child should be evaluated individually for the appropriate frequency response.

2. There was a significant improvement in the speech reception/production abilities of both groups of children during the preschool application of Verbo-tonal instruction. These results demonstrated that children with severe hearing impairments can be trained auditorily by an effective aural/oral method of instruction, such as the Verbo-tonal method. The consistent use of the appropriate amplification in conjunction with an effective instructional method may be more important to the speech and language development of a hearing-impaired child than the frequency response of the aid he uses.

3. Filtered speech testing for young deaf children provides similar information to that provided by pure-tone audiometry. Although it does not provide additional information for testing young deaf children, it is useful with older children and adults who have developed a higher level of perceptual skills.

4. The new testing measures developed during this project are useful for evaluating progress in speech production from the sub-intelligible to the intelligible level. These measures included: a similarity scale, phonetic count, and acoustic parameters. Each has certain advantages. A conventional intelligibility test is not an appropriate measure for evaluating the speech of young deaf children because most of these children need a considerable amount of training before their speech is intelligible.

ACKNOWLEDGMENT

This written Final Report reflects the effort of many persons. It includes the classroom teaching and testing of the children; the obtaining, analyzing, and interpreting of the data; and the writing, typing, and editing of the written text. The contribution of each person was extremely important. I would like to extend my deepest appreciation to all those who have contributed to this project. Without the individual efforts of each person, this study would not have been possible. I am proud to be associated with such a fine group of people.

I would also like to extend my appreciation to Drs. Robert C. Bilger, Norman P. Erber, and Charles E. Speaks, who served as consultants during the third and fourth years of this project.

APPENDIX A

THE USE OF LOW-FREQUENCY AMPLIFICATION WITH THE HEARING IMPAIRED: A REVIEW OF THE LITERATURE

The ultimate goal of educational programs for the hearing-impaired child must be the integration of the child into the normal-hearing and -speaking world (Connor, 1972). It is an illusory notion to suppose that this goal is attainable through any other means than the development in the child of speech patterns that are within the normal range of intelligibility. That most training institutions for the deaf do not achieve this goal is evidenced by the vocational and social isolation assumed by many of their graduates (Asp, 1973).

Since the hard-of-hearing or deaf child frequently has more residual hearing in the lower frequencies, efforts which emphasize the efficient use of low-frequency energy would seem to have a logical basis.

In recent years, several investigators have attempted to quantify the low-frequency hearing abilities of severely-hearing-impaired and deaf listeners. Ling (1963, 1964a, 1964b) was one of the first in the U.S. literature to suggest the value of wearable extended-range aids. He described techniques for their use and reported preliminary observations on several hearing-impaired children who had used both conventional aids and extended-range aids. He stated that the latter type of aid increased their awareness of sound and improved their voice and speech patterns. Ling (1966), in a subsequent, better-designed study, again compared the effects of conventional amplification and low-frequency amplification on the reception of speech by profoundly-deaf children. Each of two matched groups wore low-frequency aids for one week (the groups alternated the order). After each week the children used the aid they had been wearing that week to perform six live-voice listening tasks. The low-frequency aids resulted in higher test scores by both groups for detection of voiced phonemes, for identification of the number of syllables in words and phrases, for identification of stressed words in phrases, and for recognition of vowels in words. Ling concluded that the provision of maximal acoustic information is essential to rapid development of speech and language skills.

Briskey and Sinclair (1966) reported changes in the voice and speech patterns of several deaf children who wore low-frequency aids and received special auditory training. They claimed improvement in speech patterning and pitch control and cited the spectrograms of four selected cases as evidence. Briskey, Garrison, Owsley, and Sinclair (1967) subsequently evaluated voice and speech production improvement in 34 hearing-impaired children, half of whom wore conventional aids and half of whom wore low-frequency emphasis aids. They found that 11 of the 17 children using low-frequency aids improved in voice and speech quality, while only three of the 17 children using conventional aids improved. They concluded that wearable low-frequency hearing aids could be beneficial to some deaf children.

Leckie and Ling (1968), in a study with a different emphasis, investigated the phoneme-detection thresholds of 12 children with congenital hearing losses in excess of 65 dB at 500 Hz and residual hearing only for low frequencies. They were tested with a conventional response aid (300-3500 Hz) and with an experimental low-frequency response aid (80-3500 Hz). Test stimuli were the following speech sounds: /a/, /u/, /i/, /n/, and /ʃ/, recorded by both a male and a female speaker. Use of the low-frequency aid resulted in significantly-lower thresholds except for /a/.

Contrary to the above reports of successful use of low-frequency amplification, inconclusive results were obtained in an unpublished study by Hirsh and Shore (1964-1966), cited by Erber (1971b). This study compared the reception of speech by two groups of profoundly-deaf children through conventional and low-frequency emphasis aids. Each group used a different type of amplification for one year and was tested prior to the study, after three months, and after one year with both types of hearing aids. Little improvement resulted in seven listening tasks for either group over the one-year period, and no significant differences were found between groups after the period. An exception to this pattern was that low-frequency amplification resulted in superior discrimination of spoken stress patterns by both groups both before and after the experimental period. The generally-inconclusive findings were attributed in part to the influence of high levels of low-frequency ambient noise in the children's home and play environments.

The low-frequency ambient noise problem was recognized by Davis, et al. (1947) in the Harvard Report. Regarding the amplification of low frequencies, they stressed their conviction that the amplified low-pitched components of ambient noise or background speech masks the higher-pitched components of speech. Accordingly, they recommended that the frequency response be uniform, without marked resonant peaks, from 300 to 4000 Hz, with sharp cut-offs below and above this range.

Hirsh (1970) has stressed that the main reason for providing <u>any</u> kind of amplification for pre-lingual deaf children is not to increase discrimination of particular speech sounds but to yield for the child information regarding a) voice intonation, b) syllabic and verbal segmentation, c) other aspects of rhythmic structure, and d) <u>some</u> discrimination of speech sounds, "based primarily on the low-frequency sounds that are audible." Although the foregoing may be interpreted as conveying positive sentiments toward the use of low-frequency amplification, Hirsh has also asserted that "if a child responds only to frequencies below 500 Hz, we can hope only to teach him to use such residual hearing for certain patterns of rhythmic structure, of intonation, and perhaps also of stress." Clearly, Hirsh does not express a great deal of confidence in the long-term use of low-frequency amplification in the speech training of profoundly-deaf children. However, he also has stated that, although the evidence is not at all clear, it is "plausible" that the "deaf" quality of the voices of profoundly-deaf children does not appear until they begin their training in a school for the deaf where the emphasis is on visual and tactual cues.

In addition to the use of low-frequency amplification, another method of amplification has come into use in recent years which likewise attempts to make use of residual hearing in congenitally-hearing-impaired children. This method entails the use of frequency-shifting devices. Through electronic techniques, the high-frequency energy in speech sounds can be used to generate analogue signals in the lower ranges, and it has been suggested that amplification systems incorporating this principle might improve the communication ability of deaf children. Johansson (1961) was one of the pioneers in this area, developing an instrument which he called the Transposer. It selectively filtered speech energy above 4000 Hz, modulated it with a carrier frequency of 5000 Hz, and mixed the difference components below 1500 Hz with the original speech signal. The Transposer left the main vowel energy in speech unchanged and shifted the energy in the higher frequency phonemes to the low-frequency region.

Others (Wedenberg, 1961; Johansson and Sjogren, 1965; Johansson, 1966; Ling, 1968; Oeken, 1963, Ling and Druz, 1967; Hirsh, Greenwald, and Erber, 1967 [cited by Erber, 1971b]; Piminow, 1963 and 1968; Guttman and Nelson, 1968; Guttman, Levitt, and Bellefleur, 1970; Ling and Doehring, 1969; Lafon, 1967; Ling and Maretic, 1971) have reported studies using frequency-shifting devices of various types. Unlike with low-frequency amplification, however, the speech performance of children trained with these devices in general has tended not to differ significantly from that of children trained with conventional amplification systems.

Several writers have pointed out acoustic factors of various types that may operate to depreciate the perceptual (and, consequently, production) performance of hearing-impaired children who are trained with low-frequency amplification devices.

Erber (1971a) compared the abilities of children with normal hearing and children with impaired hearing to perceive speech stimuli under a range of signal-to-noise conditions. His stated goal was to define some of the S/N conditions under which provision of amplified sound can improve the reception of speech by hearing-impaired children who rely primarily upon lipreading. Using 240 "common nouns" as stimuli, he found that both his profoundly-deaf and his severely-hearing-impaired subjects required higher S/N ratios (-10 and -17 dB, respectively) for auditory detection of words than did normal-hearing children (-23 dB). He concluded that "until other corroborative studies are completed, one can only estimate that hearing-impaired children require about a 10-15 dB greater S/N ratio (i.e., about 0 to +5 dB S/N) than normal-hearing children need (i.e., about -10 dB S/N) for maximum intelligibility of speech through auditory-visual reception in low-frequency noise." He interpreted this result to mean that, although in a typical low-frequency noise classroom environment hearing-impaired children should function well auditorily (because the teacher will speak consistently close to the microphone of the auditory trainer), they will function only "marginally" outside the classroom since the S/N ratio will no longer be so favorable (because of their individually-worn hearing aids). Although Erber drew no conclusions about the use of low-frequency amplification, it seems warranted to assume that his findings would predict poor performance by hearing-impaired children if S/N conditions were anything but ideal.

Another acoustic factor that may operate negatively in the presence of low-frequency amplification is upward spread of masking, which has been found to occur in the presence of intense low-frequency sound (Bilger and Hirsh, 1956). This phenomenon, which studies have indicated occurs to a greater degree in ears with sensorineural hearing loss (Jerger, Tillman, and Peterson, 1960; Rittmanic, 1962), conceivably could be a serious limitation to the use of low-frequency amplification. Based on this evidence, Martin and Pickett (1970) designed a study to determine, for a wide range of sensorineural losses, the effects of a low-frequency masking on hearing threshold at higher frequencies. They felt that amplified environmental noise (resulting from the use of low-frequency amplification) might be intense enough to produce spread of masking. And, since much speech energy is found in the middle frequencies (French and Steinberg, 1947; Fletcher, 1953), excessive masking in this region might further reduce speech discrimination ability in impaired ears. Their study indicated that degree of loss, configuration of loss, and level of masking noise appeared to have a marked influence on upward spread of masking patterns in sensorineural subjects. They concluded that "there may be cases where the use of a hearing aid with extended low-frequency response would result in poorer aided discrimination, due to spread of masking, than would be attained with an aid having the conventional low-frequency roll-off" (Martin and Pickett, 1970).

Subsequently, Martin and Pickett (personal communication, 1972) found that the first formant (F_1) of synthetic vowels may have a low-frequency masking effect on the higher formants. For some sensorineural subjects, this masking occurred even when F_1 was presented to one ear and F_2 to the other ear, suggesting that, although in some subjects upward spread of masking occurs peripherally, in others the presence of intense low-frequency speech components somehow reduces the "apparentness" of mid- and high-frequency cues rather than causes a real shift in detection threshold. This finding appears to corroborate their findings concerning the masking effects of low-frequency noise on the aided discrimination of sensorineural subjects.

A third factor which should be taken into account in a consideration of the use of low-frequency amplification is the long-term average speech spectrum. If it can be established with assurance that low-frequency energy in speech signals does not extend significantly lower than, say, 90-100 Hz, then it would be most difficult to establish a convincing rationale for the use of an amplifier whose frequency response extends far below this limit.

French and Steinberg (1947) reported an idealized long-term average speech spectrum at one meter from the lips in a nonreverberant sound field. Both men's and women's voices were averaged in the spectrum.

Benson and Hirsh (1953) also reported measurements of the long-term average speech spectrum, presenting individual spectra for men's and women's voices. Although the ordinate scale of Benson and Hirsh's spectral graphs differs slightly from that of French and Steinberg's, both figures are similar. The three spectra have essentially the same shape except for greater energy in the 75-150 Hz octave band in the male spectrum of Benson and Hirsh caused by the lower male fundamental frequency. What is of particular importance to us, however, is not the shape of these curves, but the fact that none of them provide any specific information about the presence of speech energy below 100 Hz. Because of this fact, it seems valid to assert that it has not yet been established with certainty whether or not significant amounts of speech energy do exist below 90-100 Hz. Although the weight of evidence (through omission, however) would seem to support the contention that no significant energy exists in this region (Corliss, personal communication, 1973), the question still appears to be open.

A final crucial factor in the education of congenitally hearing-impaired children that is related to the use of low-frequency amplification is early auditory intervention. Many writers, especially in recent years, have stressed the need for the earliest-possible initiation of educational programs for these children. Wedenberg (1951) stressed the importance of early education, noting that gifted deaf children have an initial advantage in perception which the less gifted can make up through intensive training. Whetnall (1964) also stressed the importance of providing stimulation at an early age. Most recently, Connor (1972) has issued a strong plea for early educational intervention. There have been few dissenting voices on this subject, but occasionally studies have been reported which question the inherent worth of early educational intervention for hearing-impaired children. For example, Craig (1964) reported the results of a study which failed to support the hypothesis that the effects of preschool education would be evidenced by significantly-higher lipreading and reading skills favoring preschool training for deaf children. He concluded that preschool programs should be re-evaluated in terms of their goals, admission policy, and educational program. Early <u>auditory</u> intervention, however, has been the area of greatest emphasis. Hudgins (1954) stressed the importance of early auditory training, noting that, while auditory discrimination for the profoundly deaf "seems impossible," the auditory stimulus, as a supplement to vision, may be active in speech perception following auditory training. Watson (1961) claimed that it was better to begin the educational process at a young age so that the child learns to process auditory information for speech and language as well as visual information.

Myklebust (1954) has emphasized the importance of having the child receive auditory information so that verbal language and expression may develop.

> ...it is apparent clinically that children do not understand what is said to them until they have acquired a minimum of inner language. From the point of view of language development this indicates that an infant must first receive language for a certain period of time in order for the symbols to acquire their characteristic meanings. The infant then has the beginnings of inner language and he can comprehend certain spoken symbols. After hearing and comprehending these spoken symbols for another period of time, he begins to use them expressively, he begins to speak...
>
> This concept emphasizes, further, that there is a normal process of reciprocation between reception and expression of language. A child who is impeded, irrespective of cause, from receiving auditory stimuli will be reciprocally impeded in verbal language expression. Without normal language reception there cannot be normal language expression. (pp. 12-13)

To the extent that the hearing-impaired child is exposed to auditory information at an early age, the development of inner, receptive, and expressive language can take place at as "normal" a rate as possible.